THE DIFFRACTION PROPERTIES OF PLANE REFLECTION GRATINGS

by

Ross Campbell
R.C. McPhedran, B.Sc.(Hons.), University of Tasmania

A thesis submitted in fulfilment of the requirements

for the degree of

Doctor of Philosophy

in the

UNIVERSITY OF TASMANIA

HOBART

June, 1973

Except as stated herein, this thesis contains no material which has been accepted for the award of any other degree or diploma in any university. To the best of my knowledge and belief, this thesis contains no copy or paraphrase of material previously published or written by another person, except when due reference is made in the text of the thesis.

Ross McPhedran

Ross McPhedran.

CONTENTS

SUMMARY

Chapter 1 outlines the development of both ruling and holographic techniques for the production of diffraction gratings. The evolution of theoretical formulisms for describing the diffraction by gratings is also described. The experimental techniques available for testing the theories are considered, and an account is given of investigations into "anomalous" properties of gratings.

In Chapter 2, five theoretical formulations are considered in some detail. The improvements in accuracy and domain of applicability given by each new method over its predecessor are demonstrated numerically, for four of the five formulisms. (The fifth formulism does not lend itself to direct comparison with the others.) The rigorous theories are validated by comparisons of their predictions with experimental measurements.

Chapters 3 and 4 consider the theoretical properties of diffraction grating anomalies. In Chapter 3, attention is confined to gratings whose grooves are shallow compared with the wavelength of the incident radiation, and thus to S polarization anomalies. In Chapter 4, the behaviour of gratings having deeper grooves is examined, and properties of both P and S polarization anomalies are investigated.

Chapters 5 and 6 discuss the results of theoretical studies of blaze optimization. For triangular profile gratings, it is shown that, in general, groove apex angles close to 90^{o} are to be preferred, rather than the larger angles 110^{o} to 120^{o} (as recommended by G.W. Stroke). In Chapter 6, diffraction gratings having optimized sinusoidal groove profiles are shown to have blaze properties comparable with those of triangular profile gratings of the same line density.

In Chapter 7, a theoretical method for the calculation of the

profiles of holographic diffraction gratings produced in photoresist is described. This method is shown to give results in good agreement with all experimental observations of such profiles. Consideration is then given to the flexibility of arrangements for the formation of holographic gratings and an experimental system is discussed which fully utilizes this flexibility. The wide range of profiles which may be generated with such a system is demonstrated.

Chapter 8 is devoted to a study of the agreement between experimental efficiency measurements and theoretical calculations for the three most important types of grating profile - the triangular form, the same form including land, and the quasi-sinusoidal shape. Profile distortion is demonstrated to have pronounced effects on Wood anomalies. It is shown that, by appropriate utilization of the non-linear characteristics of AZ1350 photoresist, improvements of at least 70% in the spectral performance of holographic gratings can be achieved.

ACKNOWLEDGEMENTS

I wish to acknowledge the assistance given to me during the course of this research project by the following people.

Dr. M.D. Waterworth, my supervisor. He has devoted much time to discussions with me of the progress and problems of my work, and his encouragement and advice are deeply appreciated.

Mr. I.J. Wilson, my collaborator in the investigations on which Chapters 6, 7 and 8 are based. His fresh outlook and good appreciation of the practical aspects associated with our research was a great stimulus to me.

Professor R. Petit, whom I have always found to be a helpful and illuminating correspondent.

E.G. Loewen, Director of Gratings and Metrology Research, Bausch and Lomb, and M.C. Hutley, of the Division of Mechanical and Optical Metrology, National Physical Laboratory, for communicating to me the valuable experimental data on which much of Chapter 8 is based.

The overtime computer operators, Miss D. Minchin, Mrs. L. Cuffe, Mrs. M. Matthews and Mr. P. Rayner on many occasions worked long and boring shifts for me.

My fellow Ph.D. students, and staff members of the Physics and Mathematics Departments, provided me with much assistance and good cheer.

Miss Larraine Jacobs, who capably and efficiently typed this thesis.

Finally, I would like to acknowledge the contribution made to my work by my mother, who had to cope with my very irregular hours of work, and who always had confidence that I was doing something worthwhile.

During the course of my Ph.D., I was the holder of a Commonwealth Postgraduate Research Award.

CHAPTER 1

INTRODUCTION AND SURVEY OF THE FIELD

1.1 INTRODUCTION

According to Born and Wolf (1.1), a diffraction grating may be defined as "any arrangement which imposes on an incident wave a periodic variation of amplitude or phase, or both". That such objects have the ability to disperse radiation into its various wavelength components has been known for more than one hundred and fifty years. This property has been widely employed in both astronomical and chemical spectroscopy since high quality gratings became readily available.

The dispersion by a grating is governed by a fundamental equation, first derived by Joseph von Fraunhofer in the year 1821. If θ denotes the angle of incidence of radiation onto the grating (measured from its normal), and λ denotes the wavelength of the radiation, then possible directions of diffraction (θ_n) are given by

$$\sin \theta_n = \sin \theta + \frac{n\lambda}{d} \quad , \qquad (1.1)$$

where d is the period of the grating and n is an integer, positive or negative. Each value of n is associated with an order of diffraction.

A more difficult problem than that of describing grating dispersion has been that of calculating the way in which incident radiant energy is distributed among the various propagating orders formed by the grating. A consideration of various theoretical methods for the solution of this problem, and an application of one method in particular to the study of the connection between the grating characteristics and

the energy distribution function, form the first and major part of this thesis.

Until 1967, the only means of constructing diffraction gratings for use in the visible spectral region was by the use of mechanical ruling engines. The use of holographic techniques has made possible an alternative means, whose potentialities have yet to be fully exploited. The second part of this thesis is concerned with the calculation of the profile shapes of holographic diffraction gratings, and with the description of a means of producing gratings having new profile forms.

In the remainder of this chapter, the development of both ruling and holographic production techniques will be outlined. A detailed description of the evolution of adequate formulisms for describing the diffraction by gratings will be given. The experimental techniques available for testing these theories will be considered. Finally, an account will be given of investigations into the so-called "anomalous" diffraction properties of gratings.

1.2 GRATING PRODUCTION TECHNIQUES

The diffraction grating appears to have been discovered by David Rittenhouse (1.2) around 1785. Joseph von Fraunhofer (1.3) was the inventor of the ruled optical grating. Some of Fraunhofer's gratings were similar to modern spectroscopic gratings, being formed by machine ruling of grooves in metallic films deposited on a glass blank.

The improvements in the mechanical design and construction of ruling engines made by H.A. Rowland (1.4) and by A.A. Michelson (1.5) resulted in the production of the first high-resolution gratings, having resolving powers in excess of 100,000.

The last great improvement in the construction of ruling engines was made by G.R. Harrison and G.W. Stroke (1.6) in 1955, when they successfully achieved interferometric control of the M.I.T. engine. This enabled the production of gratings having much smaller systematic errors in groove position, and giving much higher resolving powers (of the order of a million) than had been obtained previously.

This survey of the development of mechanical ruling engines is necessarily brief. A much fuller exposition has been given by G.W. Stroke (1.7).

Gratings may be ruled either on spherical or flat blanks. They may have a reflecting surface, or may be used in transmission. In this thesis, we will restrict ourselves to the study of properties of the plane reflection grating, which is the most widely used type. Concave, focusing gratings are mainly used in spectral regions where the use of the separate collimating optics required by the plane grating would be a disadvantage (e.g., in the far ultraviolet region). According to Davis (1.8), "almost no transmission gratings are used in research instruments", in part because of the difficulty of making imaging transmission optics achromatic over a wide wavelength range. This leaves the plane reflection grating predominant in the near ultraviolet, visible and infrared regions.

According to Stroke (1.7), the spacing of the grooves of a high resolution grating should be accurate to better than 1/10 of the shortest wavelength of interest. Bausch and Lomb (1.9) report achieving a precision of better than 100Å. They provide gratings having up to 3600 lines/mm, with ruled widths up to 20.6 cm.

The aim of the ruling process is to achieve to a groove having a triangular profile, or (for grazing incidence gratings) a triangular

indentation in a residual of the original blank surface. (In the latter case, the remaining unruled surface is called land.) The profile obtained is governed by the shape and positioning of the ruling diamond. Ruling proceeds by a pressure-forming process, with no metal being removed (1.7). The process is complicated by the fact that the profile of the last groove ruled is appreciably deformed by the ruling of the next groove. The size of mechanically ruled gratings is limited by wear of the ruling diamond, as well as by the size of the ruling engine.

The blanks on which high resolution gratings are ruled must be of high thermal stability, and must be figured flat to within about one-twentieth of the shortest wavelength of use. The blank is generally coated with a fully-reflecting coat of aluminium, sufficiently uniform to preserve the above flatness tolerance.

Commercially available gratings are generally not prepared by ruling, but rather by replication from a ruled master grating. A detailed description of the principles of the replication process has been given by G.W. Stroke (1.7).

While research still continues into the ruling process and the improvement of ruling engines, another method for the generation of gratings has been developed. This is the holographic production technique.

For some years, it has been recognized that holograms may be viewed as a superposition of elementary diffraction gratings (1.10). It is a natural extension of this idea to contemplate the formation of diffraction gratings by holographic means.

D. Rudolph and G. Schmahl (1.11) constructed the first high-

quality holographic gratings in 1967. These gratings were made by recording the interference pattern due to two plane waves in a photoresist layer, and then etching the layer away in a solvent to leave a surface relief structure. Reflection gratings having about 500 lines/mm, and dimensions of 10 cm by 10 cm were so constructed.

A. Labeyrie and J. Flamand (1.12) were able to produce holographic gratings having up to 3000 lines/mm in sizes up to 15 cm by 11cm. They obtained diffraction efficiencies of up to 68% with an incident 5460Å polarized light beam. The level of stray light given by holographic gratings was observed to be lower than that of the best ruled gratings.

Rudolph and Schmahl (1.13) reported the construction of a holographic grating having a resolving power of 205,000, which is 90% of the Rayleigh value. They confirmed the good stray-light properties of these gratings.

The profiles of gratings constructed in this way are governed both by the intensity distribution due to the two interfering beams and the response function of the photoresist. This response function is logarithmic for the negative photoresist Resifax A used by the French firm Jobin and Yvon in the manufacture of commercial holographic gratings (1.14). Positive photoresists such as Shipley AZ1350 can be used in a linear portion of their characteristic curve (1.15). Thus both sinusoidal and logarithmic sinusoidal profiles can be obtained for holographic gratings.

Sheridon (1.16) has described a method which can furnish quasi-triangular profiles by holographic means. Instead of the two plane waves approaching the photoresist layer from the same side, as in

the arrangements of Rudolph and Schmahl and Labeyrie and Flamand, two waves approaching from opposite sides can be used. This leads to the formation of a Denisyuk hologram (1.17), with a standing wave pattern being set up in the photoresist. During the solution process, standing wave surfaces are isolated, leaving a quasi-triangular surface relief pattern on the photoresist.

Hence, by holographic methods, gratings having sinusoidal, logarithmic sinusoidal and approximately triangular profiles can be generated. (Electronmicrographs of these three profile types are given in a later section of this chapter.) In Chapter 7, a new experimental arrangement is described which will permit a great widening of the class of profile shapes able to be made holographically.

Holographic gratings having up to 3663.6 grooves/mm, and dimensions up to 16.5 cm by 32 cm are now commercially available (1.14). These give resolutions which are 80-100% of the theoretical value, and efficiencies comparable with those of mechanically ruled gratings (if the grating period is of the same order as the wavelength).

It thus appears that (even at this early stage in their development) holographic gratings can be made having sizes, line densities and diffraction efficiencies which are comparable with those of the best mechanically ruled, commercially available gratings. In applications where very low stray light is of importance (for example, in Raman spectroscopy) holographic gratings are clearly preferable.

The holographic technique also offers other advantages over the ruling process. The first is the relative simplicity and cheapness of holographic production arrangements, when compared with the complexity and great cost of modern ruling engines. Only now has

it become possible for small research groups to contemplate producing their own gratings.

The second advantage is the greater flexibility of the holographic construction process. If instead of two plane waves, two spherical waves are used to illuminate the photoresist layer, a focusing holographic grating can be made. When the blank is spherical, the concave gratings so formed offer a number of advantages over classical concave gratings, such as better stigmatism and the possibility of better blaze properties (1.18). The grating can also be constructed with one wave appropriately distorted so that, when used in a spectrograph, the grating compensates for spherical aberration in the camera mirror. A simple modification of the holographic construction apparatus permits the making of high-power Fresnel zone plates which can be used in imaging systems for soft x-rays (1.19). A fuller discussion of these matters has been given by I.J. Wilson (1.20).

A third advantage is the freedom of holographic gratings from the size limitations imposed by diamond wear on the classical ruling process. Fig(1.1) is taken from reference (1.21), and shows changes in efficiency across a grating due to diamond wear. This places an upper limit of about 80 km on the total length of groove per ruled grating. For a grating having 2000 lines/mm and a width of 20 cm, this groove length corresponds to a ruled length of 20 cm. No factors are known which limit the size of holographic gratings. The largest groove length available on such a commercial grating is close to 200 km (1.14).

The most informed judgement on the relative roles of holographic and ruled gratings would be expected from an expert in both holography and the ruling art. According to G.W. Stroke (1.22):

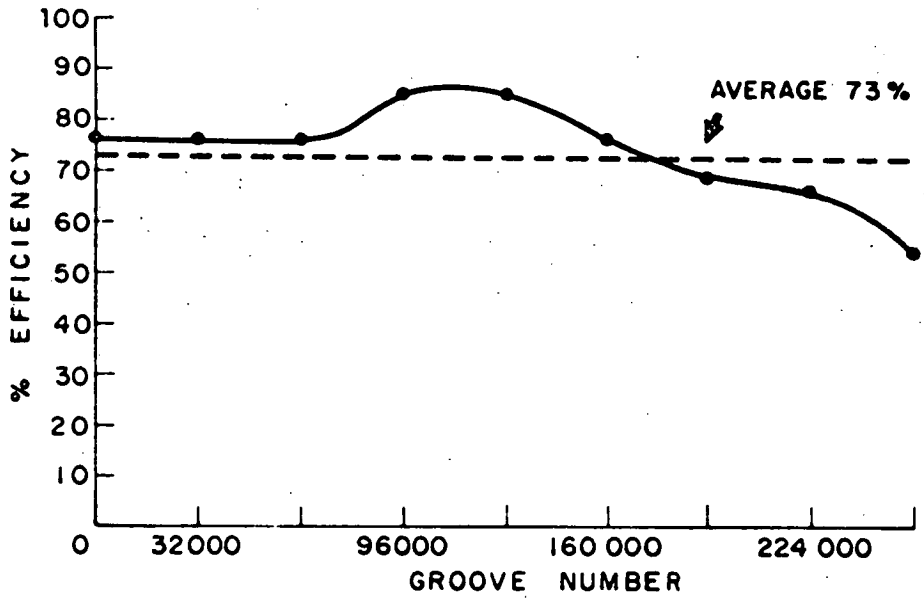


Figure (1.1). An efficiency curve in the first-order red for a grating having 261,000 grooves, showing a gradual increase with diamond wear of efficiency and its later decrease. (After Harrison et al.)

"It thus appears that the holographic gratings (usable for the visible to far-uv, and possibly for x-rays) on the one hand, and, on the other, the interferometrically servo-controlled ruled gratings (usable from the visible to the infrared and far-infrared) have come successfully to solve needs in complementary spectroscopic domains, which neither of the two grating types could have solved alone within the current state of the art".

1.3 ELEMENTARY THEORY OF HOLOGRAPHIC GRATINGS

In this section, we will consider the intensity distribution in a photoresist layer resulting from the interference of two plane waves. The period of the holographic grating formed by development of the layer will be related to the geometrical configuration of the interfering waves.

Two beams of amplitudes A_1 and A_2 propagating in the xOy plane have wave vectors making angles θ_1 and θ_2(in air) with the normal to the surface of a flat photoresist layer, as in Fig.(1.2). If the photoresist has refractive index n, then the beam angles corresponding to θ_1 and θ_2 in the layer are respectively θ_1' and θ_2', where

$$\sin\theta_1' = \frac{\sin\theta_1}{n} \quad , \quad \sin\theta_2' = \frac{\sin\theta_2}{n} \, . \qquad (1.2)$$

As is the situation in photographic emulsions (1.23), the interaction of the light waves with the photoresist is governed by their electric field strengths alone, and not by the magnetic field strengths. Hence we may describe each wave by a single scalar quantity. If the two plane waves are coherent, and have wavelengths $\lambda = 2\pi/k$ and angular frequencies ω, then we may write the associated disturbances

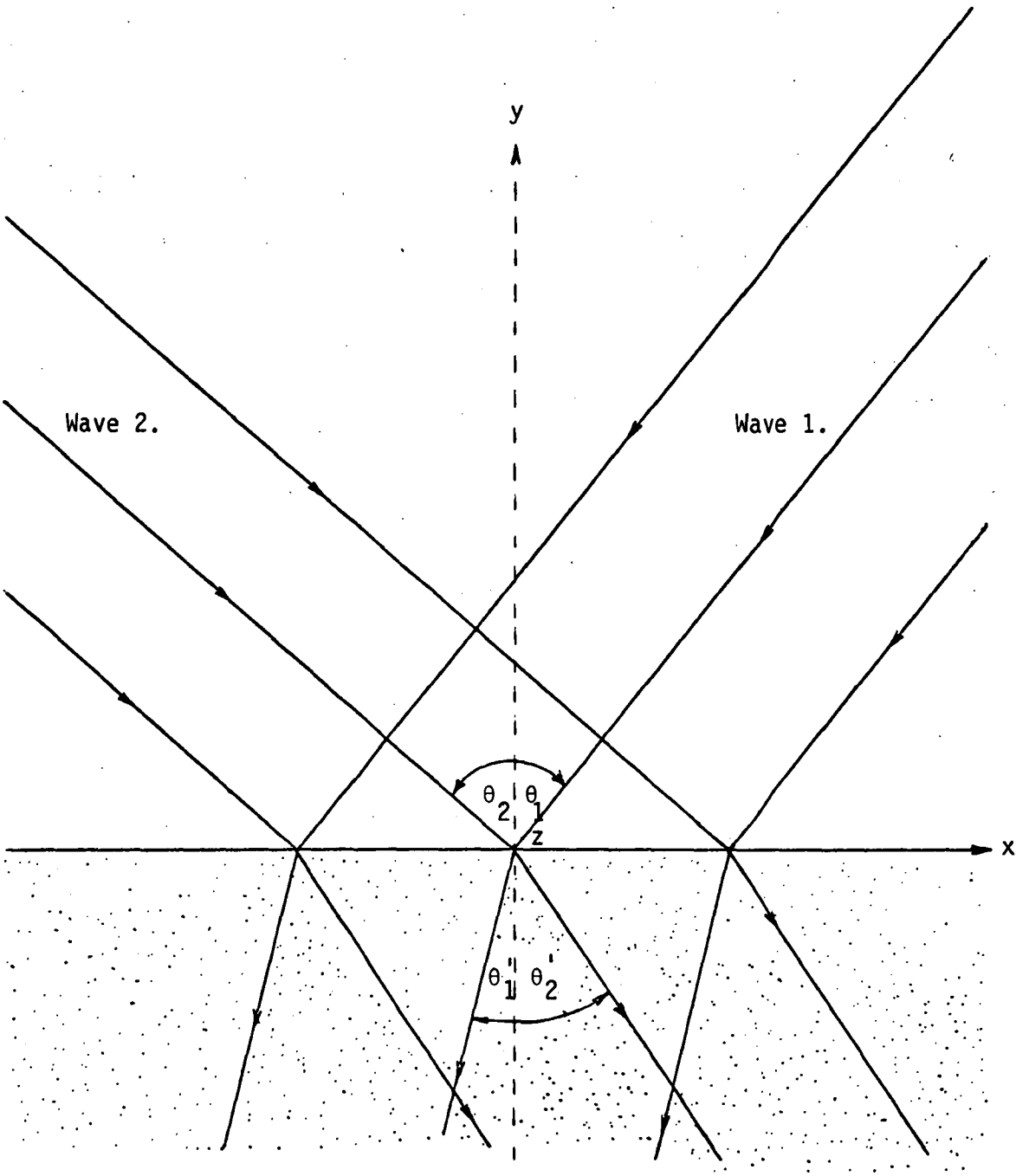


Figure (1.2). Elementary hologram formation.

in air as

$$S_1 = A_1 \cos(-kx \sin\theta_1 - ky \cos\theta_1 - \omega t),$$
$$S_2 = A_2 \cos(kx \sin\theta_2 - ky \cos\theta_2 - \omega t). \qquad (1.3)$$

The associated quantities in the photoresist are

$$S_1' = A_1' \cos(-nkx \sin\theta_1' - nky \cos\theta_1' - \omega t),$$
$$S_2' = A_2' \cos(nkx \sin\theta_2' - nky \cos\theta_2' - \omega t). \qquad (1.4)$$

The transmitted beam amplitudes A_1' and A_2' are related to A_1 and A_2 by Fresnel's equations.

The intensity of the interference pattern inside the photoresist ($I(x,y)$) is given by the time average of the square of the sum of the coherent amplitudes S_1' and S_2'. Performing the time average, we find that

$$I(x,y) = \langle (S_1' + S_2')^2 \rangle$$
$$= \frac{1}{2}(A_1'^2 + A_2'^2)\left\{1 + \frac{2A_1'A_2'}{A_1'^2 + A_2'^2} \cos\left[\frac{2\pi nx}{\lambda}(\sin\theta_1' + \sin\theta_2') + \frac{2\pi ny}{\lambda}(\cos\theta_1' - \cos\theta_2')\right]\right\}. \qquad (1.5)$$

The exposure value $E(x,y)$ associated with the intensity $I(x,y)$ is obtained by multiplying I by the exposure time T. Define now an equivalent exposure time τ by the equation

$$\tau = T \cdot \frac{1}{2}(A_1'^2 + A_2'^2). \qquad (1.6)$$

The fringe contrast factor associated with equation (1.5) is

$$c = \frac{-2A_1'A_2'}{A_1'^2 + A_2'^2}. \qquad (1.7)$$

With these definitions, and using equation (1.2), we obtain the relationship

$$E(x,y) = \tau\{1 - c \cos \frac{2\pi}{\lambda} [x(\sin \theta_1 + \sin \theta_2) + y(n\cos \theta_1' - n\cos \theta_2')]\}. \quad (1.8)$$

The exposure function E(x,y) is thus a sinusoidal function of both the spatial co-ordinates x and y. The period of the exposure function in the x direction is

$$d = \lambda/(\sin \theta_1 + \sin \theta_2). \quad (1.9)$$

The y-dependent term in equation (1.8) may be regarded as giving the relative phases of the sinusoidal, x-dependent exposure distributions.

Using equation (1.9), we obtain the desired general expression for the exposure function

$$E(x,y) = \tau\{1 - c \cos [\frac{2\pi x}{d} + \frac{2\pi n y}{\lambda} (\cos \theta_1' - \cos \theta_2')]\} . \quad (1.10)$$

Equation (1.10) and the equation (1.9) relating the period of the exposure distribution (and thus of the resultant grating) to the geometrical configuration of the two beams are of fundamental importance in the study of the holographic construction process.

The apparatuses of Labeyrie and Flamand (1.12) and Rudolph and Schmahl (1.13) used incident beams symmetrically placed about the normal to the photoresist layer. In this case ($\theta_1 = \theta_2$), the exposure distribution becomes a one-dimensional function:

$$E(x,y) \equiv E(x) = \tau\{1 - c \cos (\frac{2\pi x}{d}) \} . \quad (1.11)$$

All the x-dependent exposure distributions corresponding to various depths in the photoresist are in phase. The lines of constant exposure are the lines x = constant, parallel to the Oy axis.

The standing-wave configuration of Sheridon (1.16) uses directly opposed incident beams - i.e., $\theta_2 = \pi - \theta_1$. In this case,

$$E(x,y) = \tau \{ 1 - c \cos [\frac{2\pi x}{d} + \frac{4\pi n y}{\lambda} \cos \theta_1'] \} . \qquad (1.12)$$

The lines of constant exposure make an angle θ_1' with the Ox axis.

These lines of constant exposure will be seen in Chapter 7 to play a key role in the mathematical description of the solution process in the exposed photoresist. For the general exposure function (1.10), they have equations of the form

$$\frac{x}{d} + \frac{ny}{\lambda} (\cos \theta_1' - \cos \theta_2') = \text{constant}, \qquad (1.13)$$

or, equivalently,

$$(\sin \theta_1' + \sin \theta_2')x + y(\cos \theta_1' - \cos \theta_2') = \text{constant}. \qquad (1.14)$$

Thus, the general lines of constant exposure are rectilinear, and make an angle α with the Ox axis, where

$$\alpha = [\pi - (\theta_1' - \theta_2')]/2 \quad . \qquad (1.15)$$

1.4 THE DEVELOPMENT OF THEORIES OF THE DIFFRACTION BY GRATINGS

In this section, we will follow the development of adequate methods of calculation of the distribution of energy among the various diffracted orders of a plane reflection grating. No attempt will be made to outline the mathematical details of the various formulisms; this will be done for the more important theories in Chapter 2.

1.4.1 Scalar Methods

The scalar diffraction formulae of Kirchhoff were first applied to the case of a grating having triangular grooves by

H.A. Rowland (1.24). He was able to derive an expression for the energy distribution among the diffracted orders in closed form. This expression was used by Stamm and Whalen (1.25) to make an extensive series of calculations of diffraction efficiencies for various groove shapes and configurations of use of the grating.

Hatcher and Rohrbaugh (1.26) reported an algebraic error in Rowland's analysis, and extended it to take into account the Kirchhoff obliquity factor. They also indicated the need to include the effects of shadowing of one groove by neighbouring grooves in the calculation of efficiencies. R.P. Madden and J. Strong (1.27) derived formulae which took such shadowing effects into account. Their treatment is described more fully in Chapter 2.

The above authors used only single-scattering theories. They did not take into account the contribution to the final diffracted field from radiation which had been diffracted from one groove facet to another facet, and there re-diffracted. The importance of such multiple diffraction has been made clear by experimental efficiency measurements (1.28). Janot and Hadni (1.29) have given formulae in which secondary diffraction effects are taken into account.

All the aforementioned theories work on the hypothesis that diffraction by gratings is a scalar process - i.e., if the incident radiation is decomposed into two orthogonal polarizations, then the diffraction efficiencies for each polarization should be the same. However, it has been known since the time of Fraunhofer (1.3) that gratings can act as polarizing agents. Various experimental investigations (1.27, 1.28, 1.30, 1.31) have demonstrated that the diffraction efficiencies for light polarized parallel and perpendicular to the grating grooves can be quite different. The polarization of

diffracted orders can be as high as 95% (1.31).

The scalar theory of diffraction cannot deal adequately with gratings which exhibit substantial polarization effects. Experiment shows that rulings whose period (d) is coarse compared with the wavelength (λ) give smaller polarization effects than finer rulings. However, Cerutti-Maori and Petit(1.32) have demonstrated the inadequacy of the scalar theory for values of $\lambda|d$ as small as 0.1. Since mostmodern spectroscopic gratings are designed to be used with values of $\lambda|d$ in excess of about 0.3, the need for modification or replacement of the scalar theories is evident.

Palmer and Le Brun (1.33) have described a modified scalar formulism, in which polarization effects are taken into account by introducing different phase shifts on diffraction for radiation polarized parallel and perpendicular to the grating grooves. Secondary and tertiary diffraction effects are included in the calculation. Results furnished by this theory will be compared with experimental values and the results of an electromagnetic formulism of diffraction in Chapter 4. The electromagnetic method will be shown to be superior to the modified scalar calculation.

According to Madden and Strong (1.27):

"The scalar theory works well within its limitations but a more elegant theory is needed to explain the blaze characteristics of many gratings in use today".

1.4.2 Methods Based on the Rayleigh Expansion

In 1902, R.W. Wood (1.34) discovered experimentally certain anomalous effects manifested by optical diffraction gratings. These effects appeared when the incident light was polarized with its

electric vector perpendicular to the rulings, but not when the light was of the orthogonal polarization. (These two polarizations will be respectively referred to hereafter as S and P polarizations.) Wood anomalies will be discussed in more detail in Section (1.6).

The anomalous effects, being polarization-dependent, could not be explained within the framework of the traditional scalar theory of diffraction. However, Lord Rayleigh (1.35) was able to devise a "dynamical theory of gratings" which predicted some of the properties of Wood anomalies. He used a perturbation treatment, in which it was assumed that the depth of the grating grooves was small compared with the wavelength of the incident radiation.

It was essential to Rayleigh's analysis to know the form of the diffracted field at all points above the grating surface. The far-field expansion takes the form of a sum of outgoing plane waves of constant amplitude (B_n):

$$E^d = \sum_{n=-\infty}^{\infty} B_n \exp ik(x \sin \theta_n + y \cos \theta_n), \qquad (1.16)$$

where the angle of diffraction θ_n is given by the grating equation (1.1). If $|\sin \theta_n| \leq 1$, the order n is associated with a propagating wave. If $|\sin \theta_n| > 1$, the order is evanescent, and the wave becomes exponentially damped as the normal distance (Oy) from the grating surface increases.

Lord Rayleigh made the assumption that the plane wave expansion (1.16) is valid everywhere above the grating surface, even within the grooves. We will refer to (1.16) hereafter as the Rayleigh expansion for the diffracted field.

The various refinements of Rayleigh's perturbation

treatment which were made in order to explain in more detail the properties of Wood anomalies will be outlined in Section (1.6). These analyses all sought closed algebraic formulae for the intensities of the diffracted orders.

The advent of digital electronic computers made it no longer necessary to obtain the diffraction efficiencies in closed form. Instead, a new class of treatments was devised, in which the efficiencies were obtained by numerical solution of sets of linear equations. The necessity for the use of perturbation methods was removed.

W.C. Meecham (1.36) considered the diffraction of P polarized radiation by a perfectly conducting grating with triangular grooves. Using the Rayleigh expansion, he was able to express the boundary conditions at the grating surface as an infinite set of linear equations with the diffracted field intensities as the unknowns. The set of equations was truncated, and the unknowns were expressed as a sum of terms involving integrals of harmonic functions along the grating profile, which could be numerically evaluated.

Diffraction by perfectly conducting sinusoidal gratings was investigated by G.W. Stroke (1.37), for both P and S polarizations. Stroke's analysis was extended to deal with the case of dielectric or metallic sinusoidal gratings by P. Bousquet and R. Deleuil (1.38). In this case, two infinite systems of linear equations were derived, corresponding to the two infinite sets of unknowns, the transmitted and reflected field-component amplitudes.

The case of the diffraction by an infinitely conducting grating with triangular grooves was considered by P. Bousquet (1.39, 1.40) and also by R. Petit (1.41, 1.42). Petit's analysis is similar to that of Meecham (1.36), but is given for both P and S polarizations.

It will be described more fully in Chapter 2.

All the treatments (1.35-1.42) assume that the Rayleigh expansion (1.16) for the diffracted field is valid right on the surface of the grating grooves, where the boundary condition appropriate to the problem is applied. A physical argument first given by B.A. Lippmann (1.43) throws doubt on this assumption.

Consider the diffracted field at the point P of Fig. (1.3), located inside a grating groove. This field is caused by the current density induced on the groove surface by the incident plane wave. The current element associated with the point R can only contribute to the field at P through a wave travelling out from the groove. However, a point such as Q will give rise at P to a wave travelling into the groove. This argument makes it seem likely that the field at points actually inside a groove should be decomposed into an expansion including both outward-going and inward-going waves, with the amplitudes of each sort of wave depending on position within the groove.

This reasoning concerning the form of the diffracted field is not conclusive, as Meecham (1.36) has pointed out. Fig. (1.4) illustrates a configuration for which an incident S polarized wave has been shown (1.44) to result in a diffracted field which everywhere consists of a single outgoing plane wave. The Rayleigh expansion does hold within the grating grooves.

The first rigorous determination of the region of validity of the Rayleigh expansion was made by R. Petit and M. Cadilhac (1.45). Using a technique of analytic continuation into the complex plane, they showed that the Rayleigh expansion could not be valid in the grating grooves for P polarization and sinusoidal profiles having normalized amplitudes larger than 0.072.

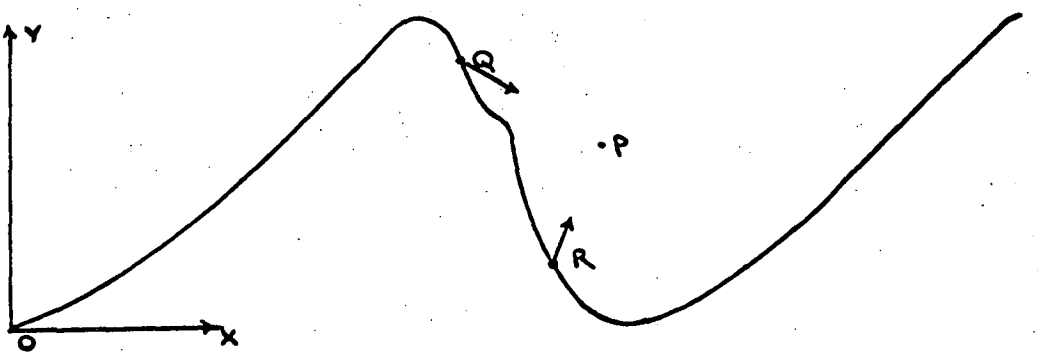


Figure (1.3). A groove of a diffraction grating. The point P will receive a wave travelling downwards from Q, and a wave travelling upwards from R.

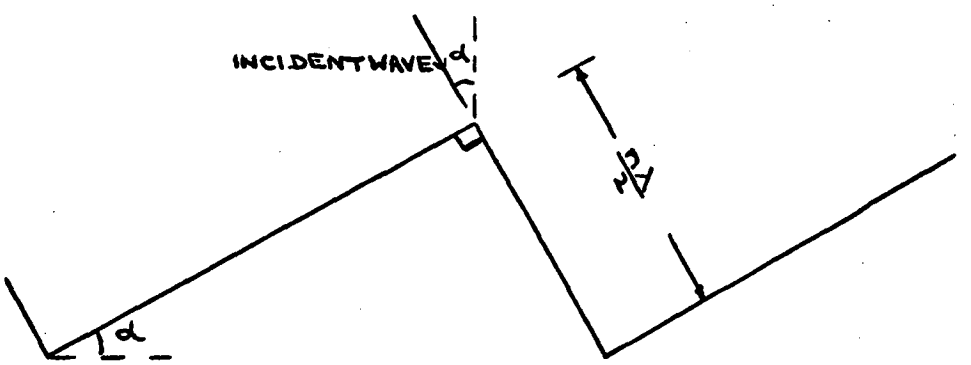


Figure (1.4). The S polarized wave shown is incident on a grating having rectangular grooves. The wave travels parallel to a groove facet whose length is an integral number of half-wavelengths. The diffracted field consists of a single plane wave, travelling in a direction opposite to that of the incident wave.

R.F. Millar (1.46, 1.47) has established that, provided the perturbation of the grating surface from a plane is sufficiently gentle, the Rayleigh expansion is always valid. He was also able to confirm the result of Petit and Cadilhac for the sinusoidal profile, and to show that Rayleigh's assumption is not in general valid if the surface of the grating is not describable by an analytic function.

M. Nevière and M. Cadilhac (1.48) have shown that (for both P and S polarizations) Rayleigh's assumption is valid only for cycloidal profiles shallower than a fixed value. It is also shown to be valid only in exceptional circumstances (e.g. the configuration of Fig. (1.4)) for triangular profiles, and any profiles with corners pointing into the vacuum half-space.

Thus, methods based on the use of the Rayleigh expansion at the grating surface are invalid for sufficiently deep grooves. They were also found to show poor accuracy for other than shallow grooves (see, for example, the convergence studies of Chapter 2). Although the methods of this section are useful in surface-diffusion studies (1.49, 1.50), a more rigorous formulism is needed for the study of diffraction by spectroscopic gratings.

1.4.3 Fourier-Series, Integral Equation Methods

The methods based on Rayleigh's assumption are all differential in character. They use the wave equation, a second order partial differential equation, in conjunction with the boundary conditions at the grating surface, to derive a system of linear equations with the field amplitudes as unknowns. We now pass on to consider formulisms in which integral equations are derived involving some unknown function, from which the diffracted field strengths may be determined.

The first such treatment was devised by R. Petit and M. Cadilhac (1.51), to deal with the diffraction of a P polarized incident wave by a grating having infinite conductivity. They were able to express the Dirichlet boundary condition on the grating surface as a Fredholm integral equation of the first kind, with the unknown function being related to the magnitude of the wave equation discontinuity at the grating surface. The Fredholm equation can be solved numerically by evaluating numerically the Fourier coefficients of the two known functions of the equation, and solving a set of linear equations for the Fourier coefficients of the unknown function (1.52, 1.53). (The details of this process will be described more fully in Chapter 2.) Petit was able to show that this Fourier series method provided results of greater accuracy than those furnished either by scalar methods or by treatments based on Rayleigh's assumption.

A similar formulation was outlined for S polarization (1.54), again leading to a Fredholm equation of the first kind. However, the use of the Fourier series method for the solution of this equation necessitated an empirical renormalization. Petit was not able to establish the legitimacy of the renormalization, but he did demonstrate that the results of the renormalized calculation were of comparable accuracy to those furnished by a Rayleigh expansion method (1.55).

Petit and Cadilhac (1.56) were able to remove the necessity for the renormalization by changing the form of the discontinuity function for the wave equation at the grating surface from a linear combination of the δ and δ' distributions to a multiple of the Heaviside function. This formulation again results in a Fredholm equation of the first kind, able to be solved by the Fourier series

method. Petit and Cadilhac demonstrated the superiority of this method over that of the renormalization. For this reason, it will be used in the convergence studies of Chapter 2.

An integral equation formulism for S polarization has been described by R.P. McClellan and G.W. Stroke (1.57), but it does not appear to have been tested numerically.

As will be shown in Chapter 2, the Fourier series methods give results of greater accuracy than those provided by Rayleigh expansion treatments, for S polarization and (more markedly) for P polarization. However, the evaluation of Fourier coefficients becomes too time consuming on a digital computer to be practicable, for profiles other than those composed of straight line segments (for which analytic expressions for the coefficients can be derived). Thus, although the formulisms of this section are quite rigorous, they did not remove the need for a method able to provide the diffraction efficiencies of general profiles.

1.4.4 The Formulation of Pavageau and Bousquet, and Similar Methods

The methods of this section again are based on the derivation of integral equations, with the induced current density on the surface of the grating being the unknown function.

The first analysis of this kind was made by J. Pavageau, R. Eido and H. Kobeissé (1.58), and resulted in a nonhomogeneous vectorial Fredholm equation of the second kind. For the case of P polarization, the vectorial equation was reduced to a scalar equation with a singular kernel (1.59). It was proposed that the integral equation be solved by an iterative method. In a final paper (1.60), J. Pavageau and J. Bousquet gave full details of the properties of the

kernel function, and of the numerical processes necessary for the calculation of the efficiencies of the various diffracted orders. The accuracy of the formulation was exemplified by numerical results of diffraction efficiencies of sinusoidal gratings, for both P and S polarizations.

D. Maystre and R. Petit (1.61) stated that for gratings of triangular profile and for S polarization the solution of the Fredholm equation by a Fourier series method results in greater accuracy than that furnished by the iterative method. However, the former method occasionally gave unsatisfactory results, for an unknown reason. The same authors later suggested (1.62) that, when the iterative method fails, it be replaced by a point-matching method.

In Chapter 2, the formulism representative of this section is that of Pavageau and Bousquet, with the integral equation being solved by the point-matching technique. The resultant method will be shown to be more accurate than the representatives of the three previous sections. Its region of convergence has not been bettered by any other method yet devised for the study of the diffraction by infinitely-conducting gratings.

Similar formulisms have been described by R. Green (1.63) and by H. Kalhor and A. Neureuther (1.64). These authors confined their attention to the study of the diffraction by gratings of triangular profile.

D. Maystre and R. Petit (1.65) have devised a method for solving the three-dimensional problem of diffraction by a grating of a wave whose wave vector is not in a plane perpendicular to that of the grating. Their analysis results in a pair of integral

equations for two components of the current density function on the surface of the grating. The solution of these equations poses no numerical problems other than those already solved for the two-dimensional case.

The same authors have shown theoretically (1.66) that a particular three-dimensional grating configuration has a transmission which is independent of wavelength. The S polarization efficiency is always 100%, while the P efficiency is larger than 80% for grating blaze angles smaller than 15^{o}. The configuration is useful for spectrometers (in principle) but not for spectrographs. It is a three-dimensional generalization of the arrangement of Fig. (1.4).

Maystre and Petit (1.67) have used the formulation of Pavageau and Bousquet to calculate efficiency curves for triangular profile gratings having 90^{o} groove vertex angles, both for a Littrow mounting and for mountings having first-order angular deviations of $\pm 15^{o}$. They showed that the effect of introducing non-zero angular deviations is not large; the form of the efficiency curves is retained and to a reasonable degree of approximation the three sets are related by angular translations. An extension of their work to groove vertex angles other than 90^{o} is given in Chapter 5.

They have also (1.68) calculated efficiency curves to establish the optimal parameters of a straight-line model of holographic grating profiles. The profiles considered were appropriate to grating formation in a logarithmic negative photoresist. The optimized profile gives a spectral performance which is practically comparable with that of the triangular profile. The blaze optimization for sinusoidal gratings (appropriate to linear photoresists) described in Chapter 6 similarly provides a good spectral performance.

1.4.5 Miscellaneous Methods

T. Itoh and R. Mittra (1.69) were able to use a wave-guide analogy in their analysis of the diffraction of a P polarized incident wave by triangular profile gratings with rectangular grooves. A modified residue calculus technique enabled the reduction of the problem to one of solving a set of linear equations. The extension of this technique to the case of non-rectangular grooves was indicated to involve some difficulty.

A differential technique for the solution of the diffraction problem for gratings with triangular grooves has recently been outlined by S. Jovicevic and S. Sesnic (1.70). They represented the field within the grooves as a superposition of wedge solutions, whose amplitudes were found and linked to the amplitudes of the plane wave field components using boundary and continuity conditions. Their results will be shown to be in poorer agreement with experimental measurements than integral equation values in Chapter 4.

P.M. Van den Berg (1.115) has applied a Green's function technique to the three-dimensional problems of the diffraction by an infinitely conducting grating of P and S polarized waves. He was able to derive integral equations of the second kind for the relevant field quantities. The equations were solved either by a Fourier series method or by using a cubic spline approximation. His numerical results for triangular and sinusoidal profiles agree with those of both Petit (1.55) and Pavageau and Bousquet (1.60).

A differential method based on conformal transformations was developed by M. Nevière, G. Cerutti-Maori and M. Cadilhac (1.71). The transformation is applied to simplify the boundary conditions, at the cost of a slight complication of the wave equation. The

transformed problem can be solved by a matrix inversion technique. The conformal transformation appropriate to a given profile is found by iteration (1.72). The accuracy of the entire method has been established by comparison with integral equation results for both sinusoidal and triangular profiles. This differential technique is of similar generality and power to the integral method of the previous section. It has also been extended to deal with the problem of three-dimensional diffraction by a grating of arbitrary profile (1.73).

1.4.6 A Finite Conductivity Theory; the Inverse Diffraction Problem

All the above formulisms had in common the assumption that the grating was made of a material having infinite conductivity, or, equivalently, that its surface reflectance was unity. Figure (1.5) shows experimental values (1.74) for the reflectance of four metals as a function of wavelength. It will be seen from the curves for the three metals (aluminium, silver and gold) commonly used in the reflecting coatings of diffraction gratings that while the identification of surface reflectance with unity is a good approximation in the infrared and millimetre-wave regions, it ceases to be possible in the visible and ultraviolet regions. Even though the development of high-reflectance coatings may make fully-reflecting gratings for the visible and near ultraviolet regions feasible in the near future, the need for a finite-conductivity theory of diffraction is clear.

Such a theory has only recently been developed. The basic formulation was first given in a study of the diffraction of a plane electromagnetic wave by a non-infinitely conducting cylinder (1.75). D. Maystre (1.76) extended the method to deal with the diffraction by a metallic grating of finite conductivity. Maxwell's equations and the

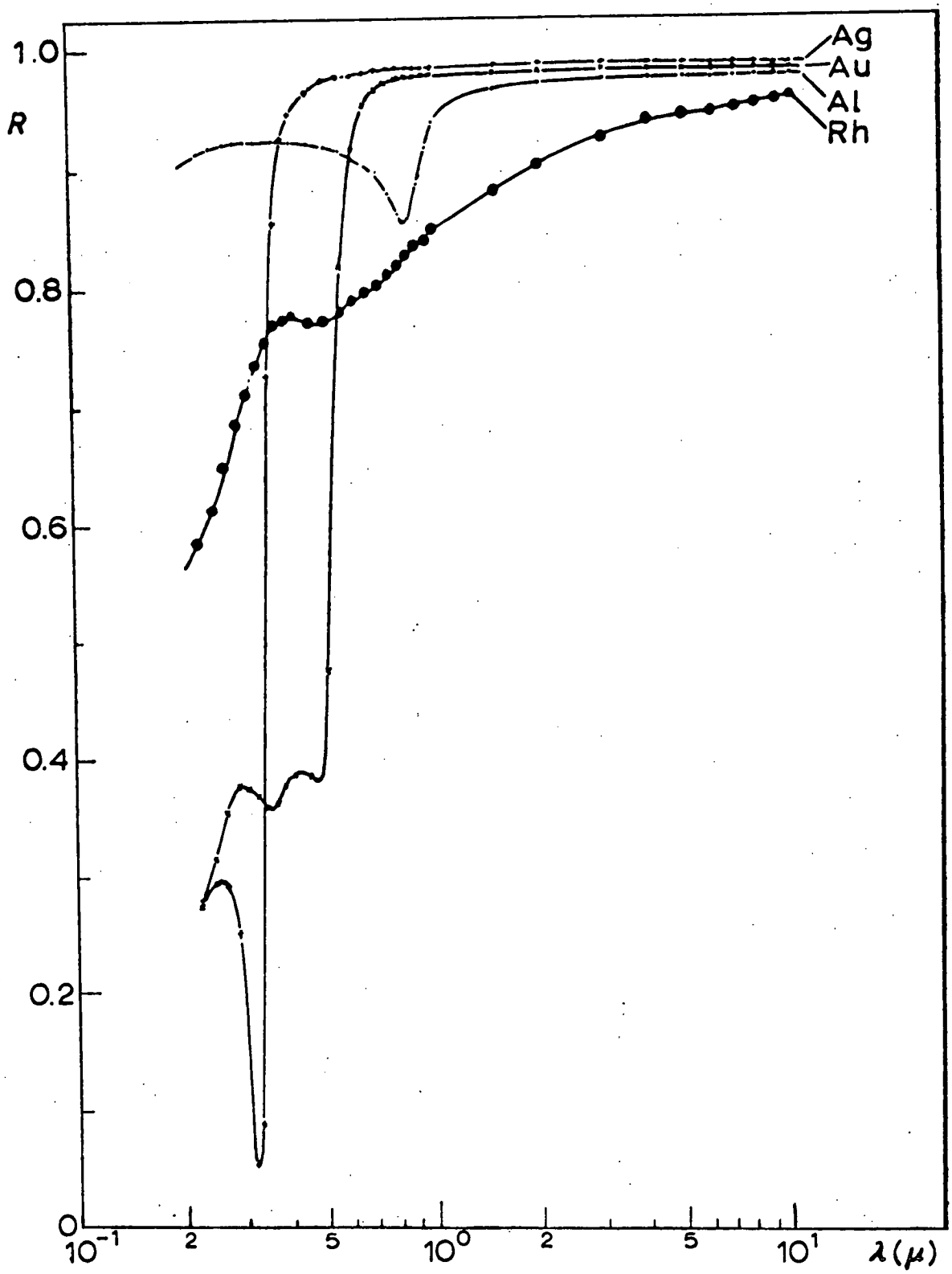


Figure (1.5). Spectral reflectance of some metals.
(After Hass and Hadley.)

boundary conditions on the surface of the grating give rise to a singular Fredholm integral equation of the first kind, with the unknown function being an equivalent surface current-density. Whereas the infinite conductivity problem required only the evaluation of single integrals, the introduction of finite conductivity leads to double integrals. The kernel function of the integral equation is much more complicated in the latter case, resulting in much larger computation times being required.

At the time of writing this thesis, only one efficiency curve for a finitely-conducting grating had been published. This is shown in Fig. (1.6), which is taken from reference (1.76). Efficiency curves for P polarized light are given for a sinusoidal grating made of an infinite conductivity material, and for one made from a material having the electrical characteristics of evaporated aluminium. The two curves have similar shapes, with the maximum deviation between them being 15%, at the peak of the P polarization blaze. They are closer together at smaller wavelengths, even though the conductivity of aluminium is lower there.

Until corresponding S polarization curves are published, it will not be possible to estimate the error which the infinite-conductivity approximation causes in efficiency values for unpolarized light in the visible region. However, the results of the infinite-conductivity theory are undoubtedly accurate for infrared and longer wavelengths. Even in the visible region, because of the computation time disadvantage of the finite conductivity theory, optimization investigations may still be primarily carried out using the simpler theory, with conductivity being taken into account only in the final stages.

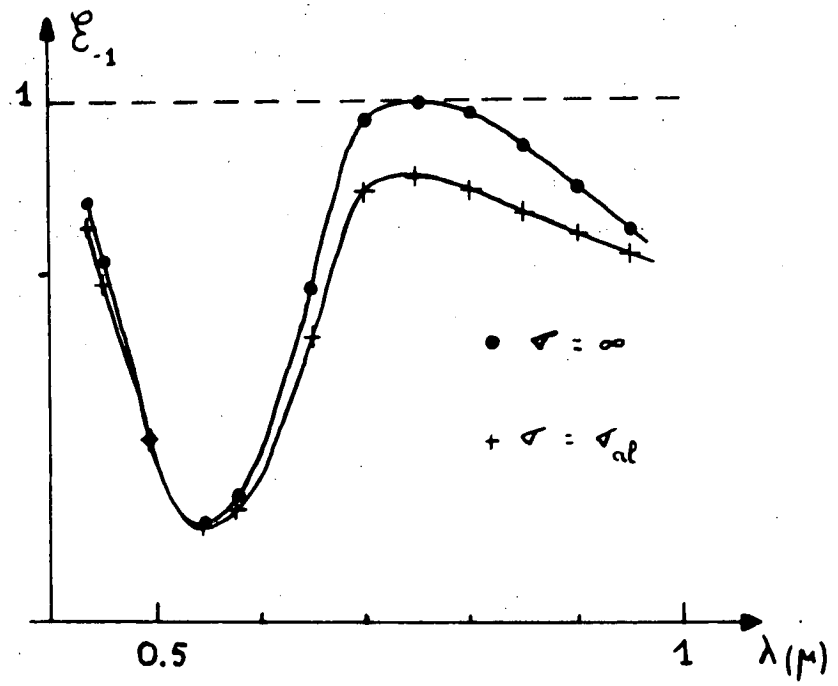


Figure (1.6). Efficiency curves for a sinusoidal grating having a ratio of groove depth to period equal to 0.4. The grating is used in a Littrow mounting in order -1, and is illuminated with P polarized light. Curves are shown both for an infinitely conducting grating, and for one made of aluminium. (After Maystre.)

Rigorous investigations of another diffraction problem have not yet begun. This is the inverse problem - given the diffraction properties of a grating, find the associated profile.

For the simple case of a grating with a period sufficiently coarse and a profile sufficiently shallow for the scalar diffraction theory to be accurate, J.A. Aas (1.77) has described a procedure for the reconstruction of the profile from an experimental diffraction spectrum. The efficiencies of the various orders are measured at a single wavelength, and the net error between the calculated and observed energy distributions is minimized numerically by the method of steepest descent.

The important problem with regard to the profiles of spectroscopic diffraction gratings is a different one. Given the efficiency in a specified order as a function of wavelength across a spectral region where a rigorous vectorial diffraction theory must be used, find the associated profile shape.

If this problem can be solved by an analytic or numerical technique, it will be a simple matter to find the grating profile which gives (say) the highest possible efficiencies over a specified wavelength interval. Profile shapes will be able to be designed to give optimum performance in a given experimental situation.

Although a direct solution of this problem will not be attempted in this thesis, two closely related matters will be studied. The first is the connection between the grating profile and the wavelength dependence of diffraction efficiency. The variation of efficiency curves with the parameters of a number of profile classes will be investigated. The second matter concerns the practical generation of that profile shape which might be determined to be optimal.

In Chapter 7, an experimental apparatus is described which will enable the formation of holographic gratings having a wide range of profile forms.

1.5. OBSERVATIONS OF PROPERTIES OF DIFFRACTION GRATINGS

In order to test a theoretical calculation of the efficiency of a diffraction grating, it must be compared with an experimental measurement of the same quantity. The efficiency measurement must be made on a grating having a known profile shape. In this section, we will briefly describe experimental techniques for the measurement of grating profiles and efficiencies.

The difficulty of determining the surface shape of a diffraction grating depends entirely on the grating period. If the profile has a period of a few millimetres, its form can be determined with ease using a normal microscope. If the period is of the order of a micron, then profile measurements must be made using an electron microscope, and consequently more difficulty is attached to obtaining results of smaller accuracy than could be achieved for millimetre-wave gratings.

The first step in the electronmicroscopic evaluation of a grating profile is the making of a replica of its surface. The replication process may result in a loss or distortion of surface details. If the replica is made in a plastic film, scale changes of the order of 2 to 10% may occur (1.78). The grating profile can then be determined by electronmicroscopic measurements on the replica.

Various techniques are used for electronmicroscopic evaluation of the replica profile. If electronmicrographs are made with the

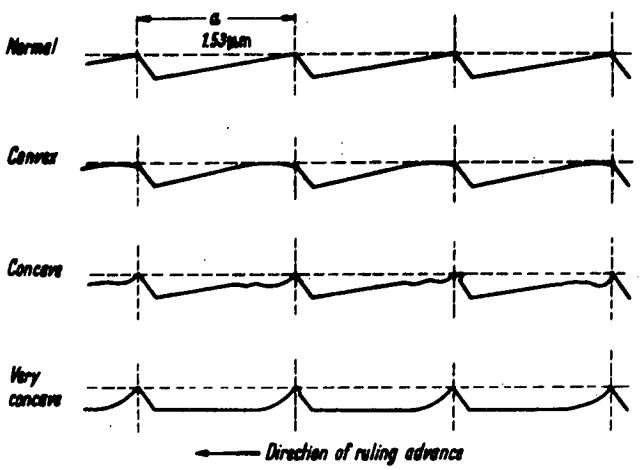


Figure (1.7). Stereoscopically-determined groove profiles of various gratings ruled in aluminium. (After Stroke.)

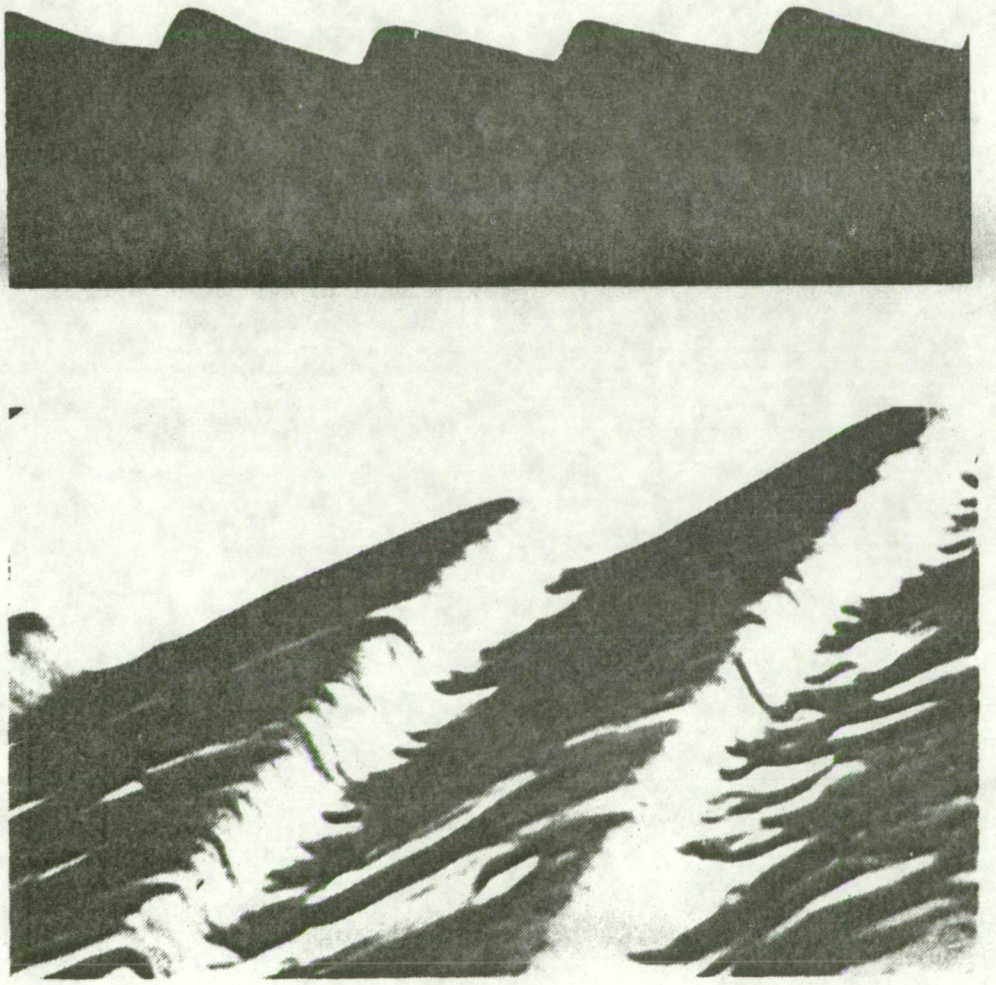


Figure (1.8). Micrographs showing the profile and the state of the surface of an echelette grating having 1230 grooves/mm. (After Bousquet et al.)

replica aligned at two different angles to the electron beam, the profile can be measured by stereoscopic methods. Figure (1.7), taken from reference (1.78) shows various groove profiles obtained by stereophotography by Pröger (1.79). The variety of groove shapes which can be produced by grating ruling is illustrated. Also evident are the considerable variations in profile which can occur in a badly ruled optical grating, even within the space of four grooves.

Figure (1.8) shows both the profile and the state of the surface of a grating having 1230 grooves/mm. The electronmicrographs were obtained by P. Bousquet, L. Capella, A. Fornier and J. Gonella (1.80), by using a replica disposed firstly parallel and then at a slight angle to the electron-beam and photographing the "shadow thrown" by the object. Despite the achievement of a good triangular profile form, substantial variations in shape occur between successive grooves. The surface micrograph shows that shape fluctuations also occur along the length of the grooves. The same shadowing technique was used to obtain the electronmicrograph shown in Figure (1.9) of the profile of a holographic grating made in a logarithmic negative photoresist (1.62). A theoretical curve corresponding to this measured profile will be given in Chapter 7.

Scanning electron microscopes can also be used to obtain micrographs of grating profiles. Three such micrographs are shown in Figure (1.10). They correspond to gratings formed holographically in a linear photoresist (1.81). The good agreement shown between the sinusoidal exposure patterns and the near-sinusoidal profiles was important in the development of an adequate theory of profile formation in holographic diffraction gratings, as outlined in Chapter 7.

Figure (1.9). The profile of a holographic grating made in a logarithmic negative photoresist. (After Maystre and Petit.)

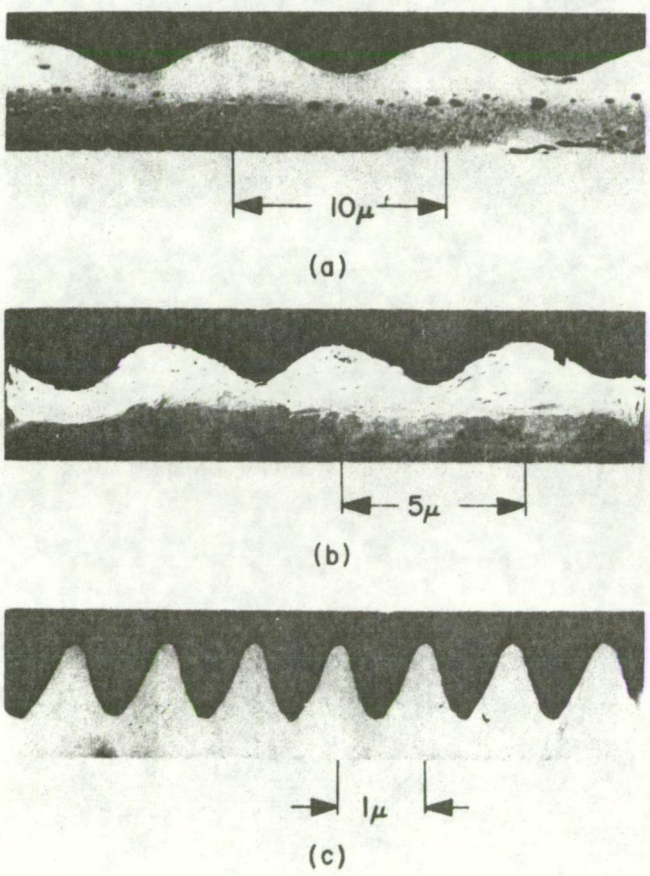


Figure (1.10). Scanning electronmicrographs of the profiles of holographic gratings made in the positive linear photoresist AZ111. (After Brandes and Curran.)

The final electronmicrograph (Fig.(1.11)) shows the quasitriangular profile obtained by N.K. Sheridon (1.16), using a photoresist exposed to a standing-wave interference pattern. Again, a corresponding theoretical profile curve will be given in Chapter 7.

An estimate of the errors inherent in using an electron-microscopic method for the measurement of diffraction grating groove geometry has been given by W.A. Anderson, G.L. Griffin, C.F. Mooney and R.S. Wiley (1.82). They determined the length of the minor facets of a grating's grooves to be $450\text{Å} \pm 65\text{Å}$. This gave a grating blaze angle of $3.3^{o} \pm 0.5^{o}$ - i.e., a measurement accuracy of ± 15%. Such considerable errors in groove shape measurement, together with the profile variations along and perpendicular to the ruling direction seen above, make difficult the comparison of theoretical efficiency values and experimental measurements made at visible wavelengths.

We turn now to a consideration of the experimental techniques used in the measurement of grating efficiency. An excellent description of efficiency testing in the visible region has been given by G.W. Stroke (1.83). He defines the measured or relative efficiency to be the ratio of the diffracted energy in a particular order at a given wavelength to the energy reflected in the same wavelength by a mirror of equivalent aperture and coated with the same surface-coating as the grating. The theoretical or absolute efficiency is the ratio of the diffracted energy at a given wavelength to the incident energy in the same wavelength. To measure relative efficiency, radiation from a broadband source is passed through a monochromator to isolate the wavelength of interest, and then is either diffracted by the grating under investigation or is reflected by the reference mirror to a

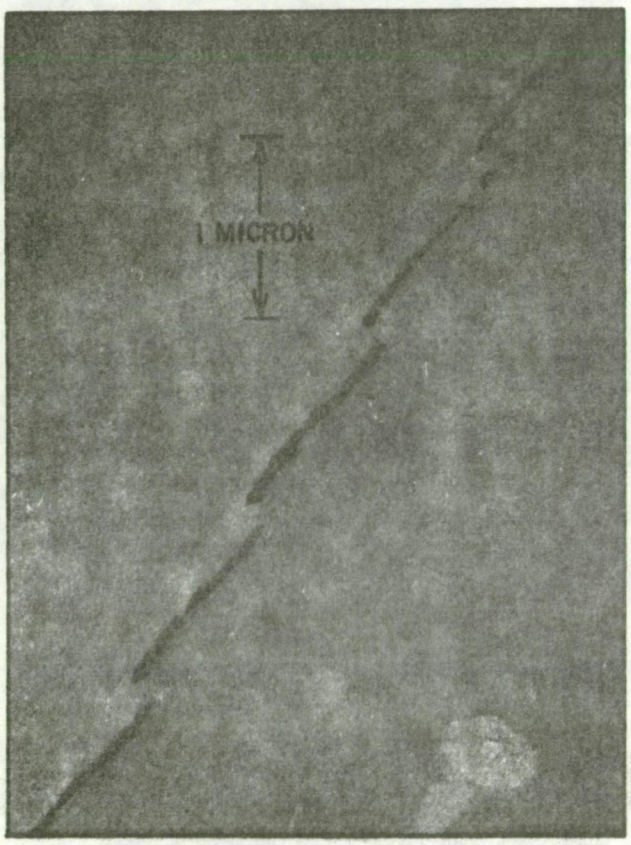


Figure (1.11). An electronmicrograph showing a quasitriangular grating profile obtained by exposing a photoresist with a standing-wave interference pattern. (After Sheridon.)

photodetector where the energy measurements are made.

Efficiency testing in the visible region provides a valuable estimate of the quality of the ruling and blazing of optical gratings. However, it was recognized as long ago as 1910 by A. Trowbridge and R.W. Wood (1.84) that investigations of the link between grating profiles and efficiencies can most advantageously be made in wavelength regions where the groove form can be determined with certainty. For this reason, their efficiency measurements were made in the infrared, rather than in the visible. More recent investigations have been carried out at even longer wavelengths, in the millimetre-wave region.

Perhaps the most sophisticated apparatus for the study of gratings with millimetre waves was constructed by R. Deleuil (1.85, 1.86). It works in the Ka band (the wavelength region from 0.75 cm to 1.13 cm). The collimating and telescope mirrors are of diameter D = 53 cm, giving a ratio of λ to D close to 1/60. The small value of this ratio ensures that the wave incident on the grating can be considered to be almost plane. The angular resolution of the apparatus is 110 minutes. An interferometric technique enables the measurement of diffraction efficiencies with a relative error smaller than 3%.

The disadvantages of working at millimetre wavelengths appear to be the limited angular resolution of the telescope systems, the difficulty of obtaining a well-collimated incident beam, and the limited number of grooves on practicable gratings (generally groove numbers are smaller than fifty). However, measurements made by C.H. Palmer and F.W. Phelps (1.87) show that the last two limitations are not likely to seriously impair the accuracy of the determination of efficiency curves. The limited angular resolution means that

measured efficiencies will be integrated over a range of diffraction angles. This can lead to errors if the efficiency is a rapidly varying function of the angle of diffraction (as will be seen in Chapter 4), but is generally not a serious defect.

The comparison of theoretical efficiency values with corresponding experimental data is only meaningful when the two can be said to correspond to the same grating profile. If the experimental measurements were made at visible or even infrared wavelengths, this can not always be guaranteed. Fig.(1.12), taken from reference (1.64), illustrates the considerable discrepancies which can exist between measurements made in two different wavelength regions on gratings having nominally the same profile. The millimetre-wave values obtained by Deleuil (1.86) are in excellent accord with the vectorial theory calculations of Kalhor and Neureuther (1.64). On the other hand, the infrared measurements of Madden and Strong (1.27) deviate substantially from Deleuil's curve. This deviation must be attributed to variations of the profile of the infrared grating from the form measured by Madden and Strong. The theoretical values provided by Madden and Strong's scalar calculations can be seen to be considerably in error.

Other measurements made by Madden and Strong and by Deleuil on triangular profile gratings will be compared with theoretical calculations in Chapter 2. Deleuil's efficiency curves showing the effects of the introduction of land into a triangular profile will be given in Chapter 8, together with corresponding theoretical curves.

The subject of the comparison of efficiency measurements made by the techniques described in this section with corresponding theoretical values will be discussed in more detail in Chapter 2.

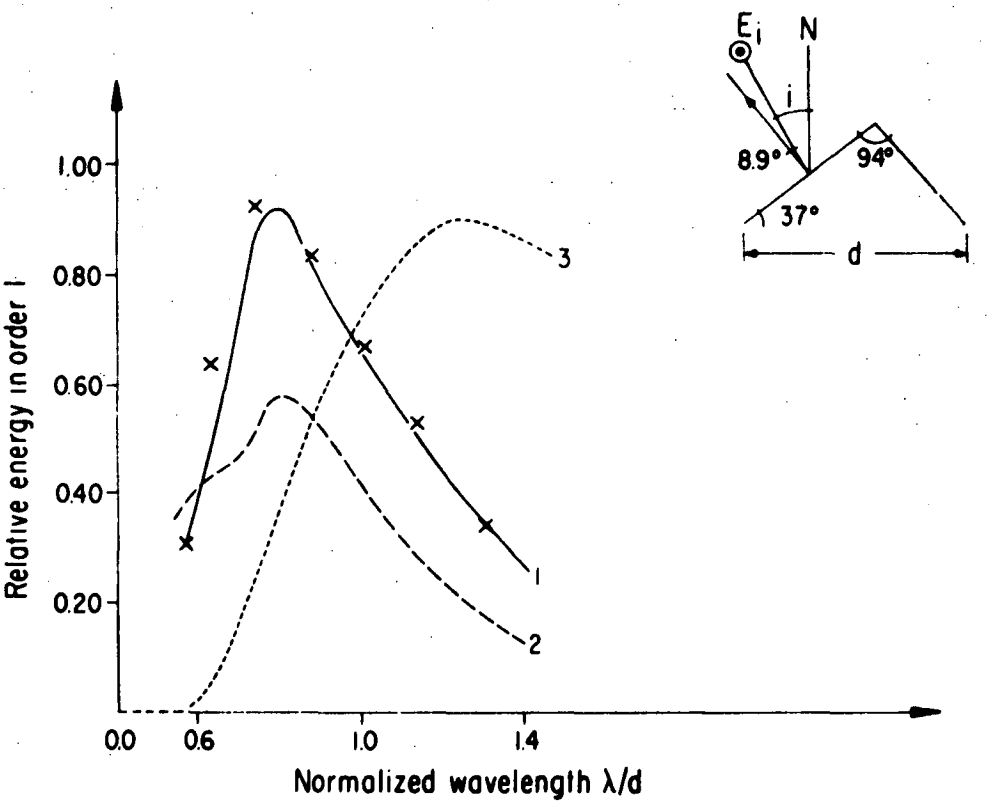


Figure (1.12). P polarization efficiency curves for an echelette grating of period d, used with a constant angular deviation of 8.9^{0} in the first order. Curves 1 and 2 correspond respectively to the millimetre-wave measurements of Deleuil, and the infrared measurements of Madden and Strong. Curve 3 was obtained using the scalar theory of Madden and Strong, and the crosses denote vectorial theory values of Kalhor and Neureuther. (After Kalhor and Neureuther.)

1.6 WOOD ANOMALIES

As mentioned above, in 1902 R.W. Wood (1.34) discovered experimentally the S polarization grating anomalies which bear his name. The anomalies took the form of very rapid intensity changes in localized regions of the diffraction spectra of a continuous source. Wood observed that the intensity dropped in one case by a factor of ten in only 6Å, so that the grating would show one of the sodium D lines and not the other. Both bright and dark anomalous intensity bands were observed. Wood attributed the occurrence of the anomalies to some factor connected with the individual groove, since he found the dark bands were of undiminished sharpness when only about 200 lines of the grating were illuminated.

It occurred to Lord Rayleigh (1.88) that the occurrence of the anomalies could be linked with the passing off of a spectrum of a higher order than the one under observation. His "dynamical theory of gratings" (1.35) used a perturbation treatment, which for S polarization diverged when an order was passing off. Lord Rayleigh was led to

"infer the probability of abnormalities in the brightness of any spectrum at the moment when one of higher order is just disappearing, abnormalities limited, however, to the case where the electric displacement is perpendicular to the ruling".

He also concluded that, when the incident light is unpolarized, the passing-off order is S polarized. The accuracy of this conclusion for shallow profilesis shown in Figure (1.13), where the S intensity curve for a sinusoidal grating is seen to come to a sharp maximum when the order passes off, whereas the P intensity curve varies smoothly there.

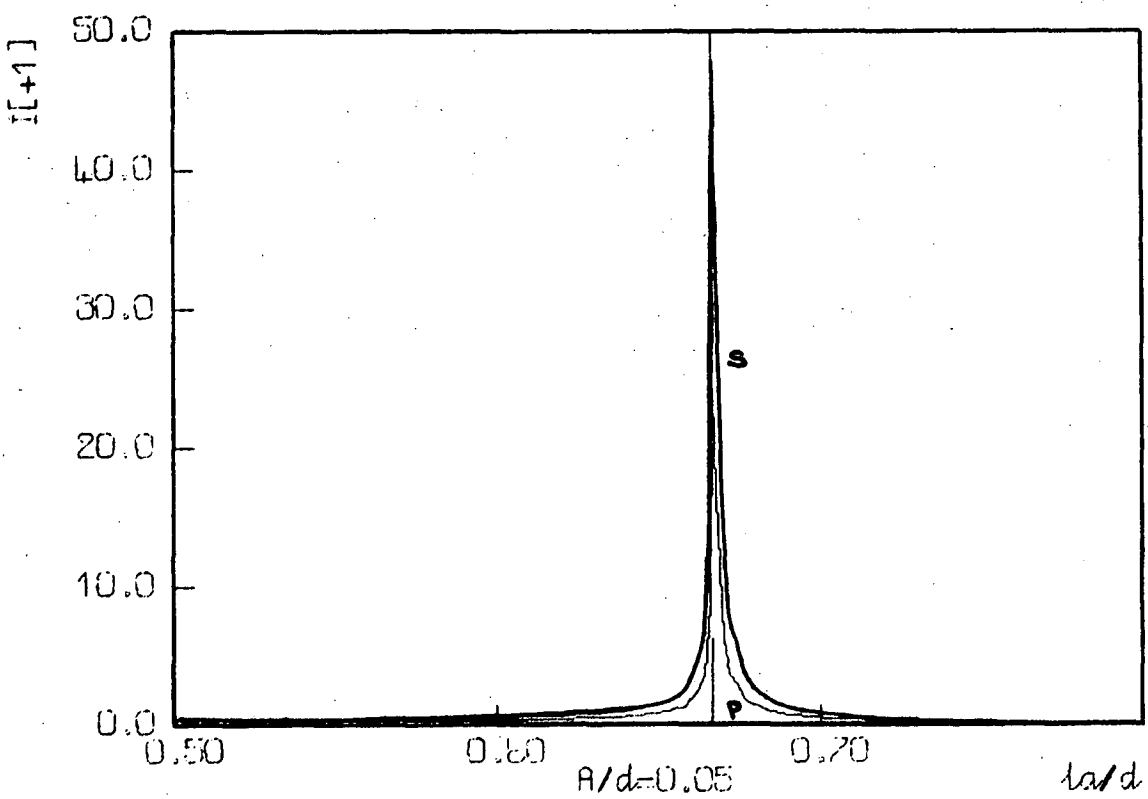


Figure (1.13). Theoretical curves showing the variations (in the order +1) of P,S and U polarization intensities as functions of normalized wavelength for a sinusoidal grating with a normalized profile depth of 0.10, used in a Littrow mounting in the order -1.

These intensity curves have been calculated using a rigorous diffraction theory.

Professor Wood (1.89) was able to confirm experimentally Lord Rayleigh's prediction that anomalous behaviour is associated with the passing off of a higher order. However, he reported the presence of anomalies in P polarized light, which could not be explained by the dynamical theory. The P anomalies were described as covering a much wider wavelength region than the corresponding S anomalies.

Lord Rayleigh's treatment (1.35) assumed the surface of the grating to be perfectly conducting. W. Voigt (1.90) extended the application of the method to surfaces of finite conductivity. U. Fano (1.91) also considered the finite conductivity case. K. Artmann (1.92) was successful in deriving a convergent series representation for the diffracted field, in the region where Rayleigh's treatment diverged. Artmann considered the maxima, but not the minima, associated with Wood anomalies.

L.R. Ingersoll (1.93) observed experimentally that the proportion of S to P polarized light in an order became larger and larger as its angle of diffraction approached 90^{o}, in agreement with Rayleigh's theory and with Fig. (1.13). He noted that the sharp sides of the anomalous minima, and not the peaks, corresponded to the Rayleigh (or passing-off) wavelength λ_R. The wavelength at the point of the minimum always exceeded λ_R. (These observations will be commented further upon in Chapter 3).

J. Strong (1.94) observed a peculiar double anomaly, which was made up of a dark band and an associated bright band. The sharp edge of the dark band corresponded to the Rayleigh wavelength.

The displacement of the bright band from the dark band was shown to depend upon the optical constants of the metal forming the grating surface. A double anomaly of this form was also described by R.W. Wood (1.95), in a paper in which the merging properties and wavelength dependence of S anomalies were discussed.

C.H. Palmer (1.96) made a systematic study of the properties of P polarization anomalies, previously only mentioned in passing by Wood (1.89, 1.95). He found that P and S anomalies have a number of properties in common. Both are associated with the passing off of higher orders. If either P or S anomalies are broad, the greatest rate of change of intensity with wavelength occurs at the Rayleigh wavelength. Both classes of anomalies are strongly affected by groove shape and by wavelength. However, the P and S anomalies in the same order may have completely different forms.

Palmer speculated that the occurrence of P anomalies is associated only with gratings whose grooves are deep compared with the wavelength. Thus, Strong and Ingersoll made their observations upon gratings with shallow grooves, and saw only S anomalies. Theories based upon the Rayleigh expansion are valid only for shallow profiles, and thus do not predict the presence of P anomalies.

Scattering theory has also been used in an attempt to explain the properties of Wood anomalies. The total diffracted field is expressed in terms of the field scattered by a single isolated element of the grating. V. Twersky (1.97, 1.98. 1.99) has used such an approach to predict for the first time the presence of anomalous minima. Theories based on the Rayleigh expansion had been used to examine anomalous maxima only. The scattering theories have been restricted to gratings having shallow profiles, and so have not proved

useful in the consideration of P anomalies. References to other applications of scattering theory in the study of diffraction by gratings may be found in Twersky's papers.

Observations on coarse gratings made by C.H. Palmer (1.100) showed that anomalies may be associated with very high grazing orders (for example, a second order anomaly was associated with the passing off of the forty-second order). He found that in some cases, strong P anomalies occurred in the absence of corresponding S anomalies, and vice versa. An observation that the anomalies changed markedly in appearance from one area to another on a grating suggests the difficulty of making experimental studies linking the form of anomalies with that of the grating profile at optical wavelengths. For this reason, most subsequent experimental investigations of anomalies have been made using millimetre waves.

Some important anomalous properties of optical gratings were observed by Stewart and Gallaway (1.101). They reported that anomalies showed a reluctance to coincide at positions where the grating equation indicated their superimposition. Instead of crossing, the anomalies were seen to repel each other, exchange identities and separate. (A theoretical interpretation of this behaviour will be given in Chapter 3). Anomalies observed on the on-blaze side of the grating were markedly different from those on the off-blaze side.

Stewart and Gallaway also derived a formula for the Rayleigh wavelength in a grating arrangement in which order n is used with a fixed angular deviation of 2α. If the order $(n+k)$ is in a grazing configuration, then

$$\lambda_R = \frac{2d}{n+2k} \left\{ \frac{1 \pm [2\sqrt{nk+k^2}/(n+2k)] \sin \alpha}{1 + [n \tan \alpha/(n+2k)]^2} \right\} , \qquad (1.17)$$

where d is the grating period. For a Littrow mounting, $\alpha = 0$ and equation (1.17) reduces to the simple expression

$$\lambda_R = 2d/(n+2k) . \qquad (1.18)$$

Observations at a wavelength of 4.2 mm made by C.H. Palmer (1.102, 1.103) have demonstrated the pronounced effect which grating surface conductivity can have on the shape of anomalies. As few as six grooves were shown to be sufficient for the exhibition of anomalous behaviour.

A fundamentally new approach to the problem of diffraction by gratings was given by A. Hessel and A.A. Oliner (1.104). Since their analysis yielded results which have proved useful in the understanding of the behaviour of grating anomalies, we will consider it in some detail in Chapter 2.

Hessel and Oliner's theory resulted in the division of grating anomalies into two classes - Wood and resonance anomalies. The passing off of a propagating spectral order causes local and rapid efficiency variations with wavelength or angle of incidence, for both P and S polarizations. These are Wood anomalies. After an order has passed off, it becomes an evanescent wave, which may resonate when its wavenumber nears that of a complex guided wave supportable by the grating. For S polarization and a shallow grating, the resonance causes a sharp intensity peak lying near the Rayleigh wavelength. (Such a peak can be seen in Figure (1.13).) As the grating groove depth increases, the intensity peak broadens and moves away

from the Rayleigh wavelength. The coupling of propagating and non-propagating waves through the boundary conditions at the grating surface means that the peaking of intensity in the resonant evanescent order causes a corresponding efficiency change in an observable order. Resonance anomalies may also be formed for P polarization, but their strength can only be appreciable if the groove depth is not shallow when compared with the wavelength. Hence, P resonance anomalies are never sharp. However, as will be seen in Chapter 4, they can have an important effect on efficiency curves.

Measurements on lamellar gratings made at millimetre wavelengths by C.H. Palmer, F.C. Evering and F.M. Nelson (1.105) showed the presence of very strong, narrow and dark S anomalies. Only weak P anomalies were found. Observations made with the same apparatus (1.87) suggest that Wood anomalies are an essentially localized phenomenon of diffraction gratings. They can be produced by as few as three grooves. As the number of grooves is increased, the diffraction anomalies become sharper, due to the increasing resolving power of the grating.

F.C. Evering (1.106) has advanced the view that the S polarization Wood anomalies of rectangular profile gratings can be considered to be due to interference between the diffracted orders of the passing-off order and those of the incident beam. Using a millimetre-wave apparatus, he has succeeded in making bright and dark anomalies appreciably brighter and darker by the introduction of an artificial passing-off order along the grating surface. He has also produced anomalies at angles of incidence where they do not normally occur. However, Evering has not considered the S polarization anomalies of complicated form which can be produced by triangular profile gratings,

and has not made artificial P polarization anomalies.

S. Fujiwara and Y. Iguchi (1.107) have applied the Rayleigh-Fano theory to the study of the diffraction by an imperfectly-conducting grating with rectangular grooves. They were able to derive approximate expressions for the intensities of the P and S polarized reflected waves around the Rayleigh wavelength. They predicted that the intensity of reflected waves should vary in parabolic fashion near the anomaly, with the radii of curvature of the parabolas being different on either side of the anomaly. This gives rise to a singularity, for both P and S polarizations. They have obtained qualitative agreement between their theoretical results and experimental measurements made in the visible region on coarse transmission gratings.

The S polarization anomalies of rectangular profile gratings have been considered theoretically by F.W. Phelps and C.H. Palmer (1.108). Their method is based on the Fraunhofer approximation, and requires experimental measurements to be made of three parameters of the theory (which are functions of the grating profile and the angle of incidence) before the efficiency values can be calculated. Experimental measurements of anomaly shapes are shown to be in reasonable agreement with theoretical predictions.

The modified scalar theory of Palmer and Le Brun (1.33) was described in section (1.4.1). Their theoretical predictions and experimental measurements will be compared with integral equation results in Chapter 4.

J.B. Breckinridge (1.31) has made extensive measurements of the polarization of diffracted light by a solar grating spectrograph. Both P and S Wood anomalies have been shown to have a marked effect on

the polarization of light by a grating. It has also been demonstrated by Breckinridge that spectrograph transmission as a function of polarization may cause large photometric and radiometric errors.

It is evident from the above survey of experimental determinations of the properties of diffraction anomalies and of theoretical attempts to explain them that much remains to be done in this field. To answer some unresolved questions, a comprehensive and rigorous investigation of the behaviour of anomalies has been made. In Chapter 3, attention will be restricted to the anomalies of shallow gratings, which are only strong in S polarized radiation. In Chapter 4, deeper profiles will be considered, and the behaviour of both P and S anomalies will be elucidated.

It should be stressed that the literature reviews contained in this chapter are not exhaustive, because of the large number of fields touched upon and the great amount of work which has been done in the past in each of these fields. A comprehensive article on diffraction gratings by G.W. Stroke (1.7) contains an excellent account of ruling engines and of the techniques which have been used for the testing of gratings. A review of the development of rigorous theories of the diffraction by gratings has been given by R. Petit and D. Maystre (1.109). This appeared in print after the completion of the bulk of this chapter. It contains some references previously inaccessible to the author (1.110-1.114).

REFERENCES

(1.1) Born, M. and Wolf, E., 1965, "Principles of Optics", 401 (Oxford: Pergamon Press Ltd.).

(1.2) Rittenhouse, D., 1786, Trans.Amer.Phil.Soc., 2, 201.

(1.3) Fraunhofer, J., 1821-1822, Denkschr.Akad.Wiss.München, 8, 1.

(1.4) Rowland, H.A., 1882, Phil.Mag., 13, 469.

(1.5) Michelson, A.A., 1912, Nature, 88, 362.

(1.6) Harrison, G.R., and Stroke, G.W., 1955, J.opt.Soc.Amer., 45, 112.

(1.7) Stroke, G.W., 1967, in "Encyclopedia of Physics", 29, 426 (Berlin: Springer-Verlag).

(1.8) Davis, S.P., 1970, "Diffraction Grating Spectrographs", 11 (New York: Holt, Rinehart and Winston).

(1.9) Bausch and Lomb, 1970, Company Publications.

(1.10) Stroke, G.W., 1966, "An Introduction to Coherent Optics and Holography", 8(New York: Academic Press).

(1.11) Schmahl, G. and Rudolph, D., 1968, Mitt.Astr.Ges., 24, 41.

(1.12) Labeyrie, A. and Flamand, J., 1969, Optics Communications, 1, 5.

(1.13) Rudolph, D. and Schmahl, G., 1970, Optik, 30, 475.

(1.14) Jobin and Yvon, Company Publication.

(1.15) Beesley, M.J. and Castledine, J.G., 1970, Appl.Optics, 9, 2720.

(1.16) Sheridon, N.K., 1968, Appl.Phys.Letters, 12, 316.

(1.17) Denisyuk, Y.N., 1962, Sov.Phys.-Doklady, 7, 543.

(1.18) Cordelle, J., Flamand, J., Pieuchard, G. and Labeyrie, A., 1970, in "Optical Instruments and Techniques" (Newcastle upon Tyne: Oriel).

(1.19) Schmahl, G. and Rudolph, D., 1969, Optik, 29, 577.

(1.20) I.J. Wilson, 1972, unpublished Honours Thesis (University of Tasmania, Physics Department).

(1.21) Harrison, G.R., Thompson, S.W., Kazukonis, H., and Connell, J.R., 1972, J.opt.Soc.Amer., 62, 751.

(1.22) Stroke, G.W., and Johnson, A.E., 1970, in "Optical Instruments and Techniques" (Newcastle upon Tyne: Oriel).

(1.23) Born, M. and Wolf, E., 1965, "Principles of Optics", 280 (Oxford: Pergamon Press Ltd.).

(1.24) Rowland, H.A., 1893, Phil.Mag., 35, 397.

(1.25) Stamm, R.F., and Whalen, J.J., 1946, J.opt.Soc.Amer., 36,2.

(1.26) Hatcher, R.D., and Rohrbaugh, J.H., 1956, J.opt.Soc.Amer., 46,104.

(1.27) Madden, R.P., and Strong, J., 1958, in "Concepts of Classical Optics", 597 (San Francisco: W.H. Freeman and Company).

(1.28) Munier, J., Claudel, J., Décamps, E., and Hadni, A., 1962, Rev. d'Opt., 41, 245.

(1.29) Janot, C., and Hadni, A., 1962, J.Physique Rad., 23, 152.

(1.30) Rohrbaugh, J., Pine, C., Zoellner, W., and Hatcher, R., 1958, J.opt.Soc.Amer., 48, 710.

(1.31) Breckinridge, J.B., 1971, Appl.Optics, 10, 286.

(1.32) Cerutti-Maori, G., and Petit, R., 1970, Nouv.Rev.d'Opt.appl., 1, 321.

(1.33) Palmer, C.H., and Le Brun, H., 1972, Appl. Optics, 11, 907.

(1.34) Wood, R.W., 1902, Phil.Mag., 4, 396.

(1.35) Rayleigh, Lord, 1907, Proc.Roy.Soc.A, 79, 399.

(1.36) Meecham, W.C., 1956, J.Appl.Phys., 27, 361.

(1.37) Stroke, G.W., 1960, Rev.d'Opt., 39, 291.

(1.38) Bousquet, P., and Deleuil, R., 1963, C.R. Acad.Sci.Paris, 256,1461.

(1.39) Bousquet, P., 1963, C.R. Acad.Sci.Paris, 256, 3422.

(1.40) Bousquet, P., 1963, C.R. Acad.Sci.Paris, 257, 80.

(1.41) Petit, R., 1963, C.R. Acad.Sci.Paris, 257, 2018.

(1.42) Petit, R., 1963, Rev.d'Opt., 42, 263.

(1.43) Lippmann, B.A., 1953, J.opt.Soc.Amer., 43, 408.

(1.44) Maréchal, A., and Stroke, G.W., 1959, C.R. Acad.Sci.Paris, 249, 2042.

(1.45) Petit, R., and Cadilhac, M., 1966, C.R.Acad.Sci.Paris, 262, 468.

(1.46) Millar, R.F., 1969, Proc.Camb.Phil.Soc., 65, 773.

(1.47) Millar, R.F., 1971, Proc.Camb.Phil.Soc., 69, 217.

(1.48) Nevière, M., and Cadilhac, M., 1970, Opt.Commun., 2, 235.

(1.49) Blakely, J.M., and Olson, D.L., 1968, J.Appl.Phys., 39, 3476.

(1.50) Bonzel, H.P., and Gjostein, N.A., 1968, J.Appl.Phys., 39, 3480.

(1.51) Petit, R., and Cadilhac, M., 1964, C.R.Acad.Sci.Paris, 259, 2077.

(1.52) Petit, R., 1965, C.R.Acad.Sci.Paris, 260, 4454.

(1.53) Petit, R., 1965, Appl.Optics, 4, 1551.

(1.54) Petit, R., 1965, C.R.Acad.Sci.Paris, 261, 4677.

(1.55) Petit, R., 1966, Rev.d'Opt., 45, 249.

(1.56) Petit, R., and Cadilhac, M., 1967, C.R.Acad.Sci.Paris, 264, 1441.

(1.57) McClellan, R.P., and Stroke, G.W., 1966, J.Math.Phys., 45, 383.

(1.58) Pavageau, J., Eido, R., Kobeissé, H., 1967, C.R.Acad.Sci.Paris, 264, 424.

(1.59) Pavageau, J., and Bousquet, J., 1969, C.R.Acad.Sci.Paris, 268, 776.

(1.60) Pavageau, J., and Bousquet, J., 1970, Optica Acta, 17, 469.

(1.61) Maystre, D., and Petit, R., 1970, C.R.Acad.Sci.Paris, 271, 400.

(1.62) Maystre, D., and Petit, R., 1970, Opt.Commun., 2, 309.

(1.63) Green, R., 1970, I.E.E.E.Trans. MTT-18, 313.

(1.64) Kalhor, H., and Neureuther, A., 1971, J.opt.Soc.Amer., 61, 43.

(1.65) Maystre, D., and Petit, R., 1971, Opt.Commun., 4, 97.

(1.66) Maystre, D., and Petit, R., 1972, Opt.Commun., 5, 35.

(1.67) Maystre, D., and Petit, R., 1971, Nouv.Rev.d'Opt.appl., 2, 115.

(1.68) Maystre, D., and Petit, R., 1971, Opt.Commun., 4, 25.

(1.69) Itoh, T., and Mittra, R., 1969, I.E.E.E.Trans. MTT-17, 319.

(1.70) Jovicevic, S., and Sesnic, S., 1972, J.opt.Soc.Amer., 62, 865.

(1.71) Nevière, M., Cerutti-Maori, G., and Cadilhac, M., 1971, Opt.Commun., 3, 48.

(1.72) Nevière, M., and Cadilhac, M., 1971, Opt.Commun., 3, 379.

(1.73) Nevière, M., and Cadilhac, M., 1971, Opt.Commun., 4, 13.

(1.74) Hass, G., and Hadley, L., 1963, in "American Institute of Physics Handbook",6(New York: McGraw-Hill).

(1.75) Maystre, D., and Vincent, P., 1972, Opt.Commun., 5, 327.

(1.76) Maystre, D., 1972, Opt.Commun., 6, 50.

(1.77) Aas, J.A., 1972, Appl.Optics, 11, 1579.

(1.78) Stroke, G.W., 1967, in "Encyclopedia of Physics", 29, 657 (Berlin: Springer-Verlag).

(1.79) Pröger, H.J., 1961, Exp.Techn.der Phys., 9, 236.

(1.80) Bousquet, P., Capella, L., Fornier, A., and Gonella, J., 1969, Appl.Optics, 8, 1229.

(1.81) Brandes, R.G., and Curran, R.K., 1971, Appl.Optics, 10, 2101.

(1.82) Anderson, W.A., Griffin, G.L., Mooney, C.F., and Wiley, R.S., 1965, Appl.Optics, 4, 999.

(1.83) Stroke, G.W., 1967, in "Encylcopedia of Physics", 29, 644 (Berlin: Springer-Verlag).

(1.84) Trowbridge, A., and Wood, R.W., 1910, Phil.Mag., 20, 886.

(1.85) Deleuil, R., 1966, C.R.Acad.Sci.Paris, 262, 1676.

(1.86) Deleuil, R., 1969, Optica Acta, 16, 23.

(1.87) Palmer, C.H., and Phelps, F.W., 1968, J.opt.Soc.Amer., 58, 1184.

(1.88) Rayleigh, Lord, 1907, Phil.Mag., 14, 60.

(1.89) Wood, R.W., 1912, Phil.Mag., 23, 310.

(1.90) Voigt, W., 1911, Göttinger Nachr., 40.

(1.91) Fano, U., 1938, Ann.Phys., 32, 393.

(1.92) Artmann, K., 1942, Z.Phys., 119, 529.

(1.93) Ingersoll, L.R., 1921, Phys.Rev., 17, 493.

(1.94) Strong, J., 1936, Phys.Rev., 49, 291.

(1.95) Wood, R.W., 1935, Phys.Rev., 48, 928.

(1.96) Palmer, C.H., 1952, J.opt.Soc.Amer., 42, 269.

(1.97) Twersky, V., 1952, J.Appl.Phys., 23, 1099.

(1.98) Twersky, V., 1962, J.opt.Soc.Amer., 52, 145.

(1.99) Twersky, V., 1962, I.R.E. Trans. AP-10, 737.

(1.100) Palmer, C.H., 1956, J.opt.Soc.Amer., 46, 50.

(1.101) Stewart, J.E., and Gallaway, W.S., 1962, Appl.Optics, 1, 421.

(1.102) Palmer, C.H., 1963, from "A Symposium on Electromagnetic Warfare", the Johns Hopkins University, Baltimore, Maryland.

(1.103) Palmer, C.H., 1964, "Symposium on Quasi-Optics" (Brooklyn: Polytechnic Press).

(1.104) Hessel, A., and Oliner, A.A., 1965, Appl.Optics, 4, 1275.

(1.105) Palmer, C.H., Evering, F.C., and Nelson, F.M., 1965, Appl.Optics, 4, 1271.

(1.106) Evering, F.C., 1966, Appl.Optics, 5, 1313.

(1.107) Fujiwara, S., and Iguchi, Y., 1968, J.opt.Soc.Amer., 58, 361.

(1.108) Phelps, F.W., and Palmer, C.H., 1969, J.opt.Soc.Amer., 59, 812.

(1.109) Petit, R., and Maystre, D., 1972, Rev.de Phys.appliquée, 7, 427.

(1.110) Van den Berg, P.M., 1971, "Rigorous diffraction theory of optical reflection and transmission gratings", thesis (Rotterdam: Bronder-offset N.V.).

(1.111) Bolomey, J.C., 1971, thesis number C.N.R.S.: A.O. 5604 (Paris).

(1.112) Maystre, D., and Petit, R., "Diffraction by an infinitely conducting lamellar grating", Opt.Commun., in press.

(1.113) Cerutti-Maori, G., 1970, thesis (Paris).

(1.114) Nevière, M., Cadilhac, M., and Petit, R., 1971, International symposium, Tbilissi (Moscow: Editions Nauka).

(1.115) Van den Berg, P.M., 1971, Appl.Sci.Res., 24, 261.

CHAPTER 2

THEORIES OF THE DIFFRACTION BY GRATINGS

2.1. INTRODUCTION

In the previous chapter, an account was given of the development of adequate methods of calculation of the distribution of energy among the diffracted orders formed by a plane reflection grating. No detailed descriptions were given of the various mathematical formulisms which enable the numerical evaluation of the energy distribution.

In this chapter, we will outline the theoretical methods used in what are felt to be five of the most important approaches to the study of the diffraction by gratings. The improvements in accuracy and domain of applicability given by each new method over its predecessor will be demonstrated numerically, for four of the five formulisms. (The fifth, that of Hessel and Oliner, is formulated in such a way that its results cannot be compared directly with those provided by the other four methods.) Finally, the rigorous theories will be validated by comparisons of their predictions with experimental measurements of diffraction efficiencies.

2.2 A SCALAR THEORY

The representative of the group of scalar theories of grating diffraction which we will consider in this section is the method of R.P. Madden and J. Strong (2.1). This is the most complete single-scattering scalar theory.

The analysis starts from the differential form of Kirchhoff's formula. In Figure (2.1), let Q' be the source of a scalar spherical

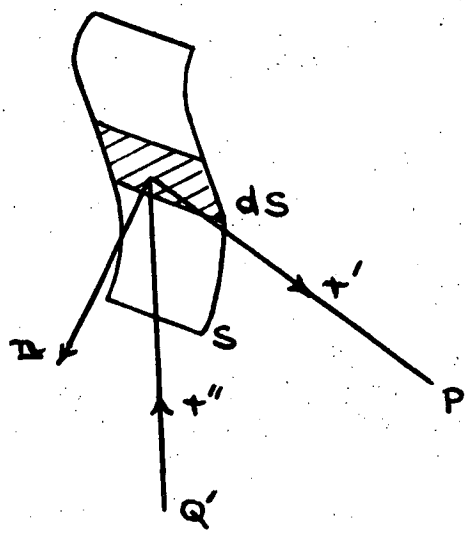


Figure (2.1). Definitions associated with the treatment of the diffraction by a surface element dS of a scalar spherical wave emanating from a point Q'. The observation point is at P. (After Madden and Strong.)

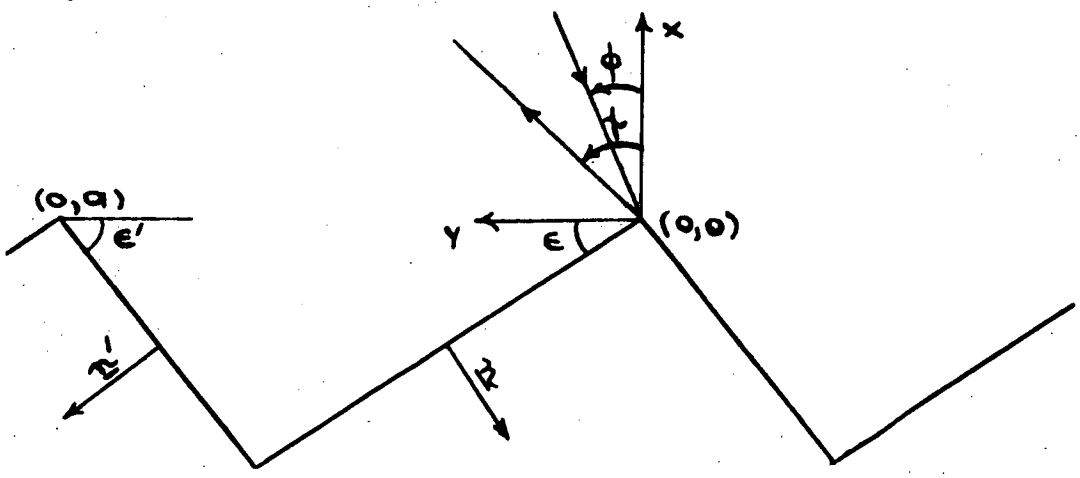


Figure (2.2). Definitions associated with the treatment of the diffraction by a grating having triangular grooves (with facet angles ε and ε'). The directions of the incident and diffracted rays are respectively specified by the angles ϕ and ψ. (After Madden and Strong.)

wave having amplitude A, wavelength $\lambda = 2\pi/k$ and angular frequency ω. An element dS of the grating surface is located at a distance r" from Q', and at a distance r' from an observation point P. The diffracted scalar field at P associated with dS is then

$$dE_P = \frac{jA}{2\lambda r'' r'} \exp[j\omega(t - \frac{r'+r''}{c})] \cdot [-\cos(n,r'')-\cos(n,r')]dS \quad (2.1)$$

where $\underline{n}$ denotes the unit normal vector to dS. The Kirchhoff's formula (2.1) can only be expected to yield accurate results if it is integrated over an aperture large compared with the wavelength (2.2, 2.3). If we introduce a Cartesian coordinate system Oxyz, with the Oz axis along the grating grooves and the Ox and Oy axes in the plane of the groove section (as shown in Figure (2.2)), and specify the incident and diffracted rays respectively by direction cosines (α, β, γ) and $(\alpha', \beta', \gamma')$, then equation (2.1) gives when integrated over the whole grating surface

$$E_P = \frac{jE_0}{2\lambda r'} \exp[j\omega(t - \frac{R+R'}{c})] \sum_{n=0}^{N-1} \exp[jkna(\beta' - \beta)] \cdot \int_{z=-\frac{h}{2}}^{\frac{h}{2}} \exp[jk(\gamma' - \gamma)z]dz \int_{\text{one groove}} [\cos(n,r)-\cos(n,r')] \cdot \exp jk[(\alpha' - \alpha)x + (\beta' - \beta)y] d\ell \; . \quad (2.2)$$

Here R and R' are respectively the distances from O to Q' and to P, and $E_0 = A/R$. $d\ell$ denotes an element of length along the groove surface, and the grating has N grooves of length h and period a. In the derivation of equation (2.2), the Fraunhofer assumption that both the source and observation points are distant from the grating surface has been made.

We now suppose that the grating has grooves of triangular form, with facet angles ε and ε', as shown in Figure (2.2). Let $c = \tan \varepsilon$, and $c' = \tan \varepsilon'$. If ϕ characterizes the direction of incidence, and ψ that of diffraction, then

$$\alpha' - \alpha = \cos \phi + \cos \psi = \rho,$$

$$\beta' - \beta = \sin \psi + \sin \phi = \mu.$$

In the evaluation of the second integral of equation (2.2), two cases arise. In the first case, the angles of incidence and diffraction are such that neither groove is in the geometrical shadow of the other for either the incident or the diffracted ray. This situation occurs when

$$(-90^{o} + \varepsilon) \leq \phi, \psi \leq (90^{o} - \varepsilon') \ . \qquad (2.3)$$

The groove integral then takes place along the total length of each facet.

The grating efficiency is defined to be the fraction of incident monochromatic flux which is diffracted into a given order of the grating (2.1). The diffracted flux is obtained by integrating the squared modulus of equation (2.2) over the exit pupil. If we define

$$x = \rho/\mu \ , \qquad (2.4)$$

and take the grating to be the aperture stop in its illuminating system, then the efficiency of the grating in the absence of shadowing is

$$E(m) = \frac{1}{\cos \phi \cos \psi} \frac{\lambda^2 (c+c')^2}{4\pi^2 a^2} \left\{ \frac{(1+x^2)}{(1-cx)(1+c'x)} \sin \left[\frac{m\pi c'}{c+c'} (1-cx) \right] \right\}^2 . \qquad (2.5)$$

Equation (2.5) gives the diffraction efficiency in order m of the

grating, provided that m is non-zero. If m = 0, μ = 0 and x becomes undefined. The zero-order efficiency is given by

$$E(0) = \left\{ \sin\left[\frac{2\pi c'c\ a\ \cos\phi}{\lambda(c+c')}\right] \Big/ \left[\frac{2\pi c'c\ a\ \cos\phi}{\lambda(c+c')}\right]\right\}^2 . \quad (2.6)$$

Consider now the case when shadowing takes place - i.e.,

$$\phi \text{ or } \psi > (90^0 - \varepsilon') \quad \text{or} \quad \psi < \varepsilon - 90^0 \quad (2.7)$$

A particular example of this case ($\phi > 90^0-\varepsilon'$) is shown in Figure (2.3). In the geometrical optics view, the element ABC of the groove is shadowed, and only the element CD can contribute to the diffracted field. This shadowing effect can be taken into account by the introduction of a foreshortening factor f. If ϕ or ψ exceeds $90^0-\varepsilon'$, and θ denotes the maximum of the two angles, then

$$f = \frac{\sin(\varepsilon+\varepsilon')}{\sin\varepsilon'} \left[\cos\varepsilon - \sin\varepsilon\tan(\theta-\varepsilon)\right] . \quad (2.8)$$

The efficiency in order m is given by

$$E(m) = \frac{\lambda^2}{4\pi^2a^2\cos\phi\cos\psi} \left\{\frac{(x+c)}{(1-cx)} \sin\left[\frac{m\pi c'f}{c+c'}(1-cx)\right]\right\}^2 \quad (2.9)$$

if $m \neq 0$. Also,

$$E(0) = \left\{\sin\left[\frac{2\pi cc'\ a\ f\ \cos\phi}{(c+c')\lambda}\right] \Big/ \frac{2\pi ac\ \cos\phi}{\lambda}\right\}^2 . \quad (2.10)$$

The formulae appropriate to the case when $\psi < (\varepsilon-90^0)$ may be obtained from the preceding three equations if the following substitutions are made:

$$\varepsilon \to \varepsilon',\ \varepsilon' \to \varepsilon,\ \phi \to (-\psi),\ \psi \to (-\phi). \quad (2.11)$$

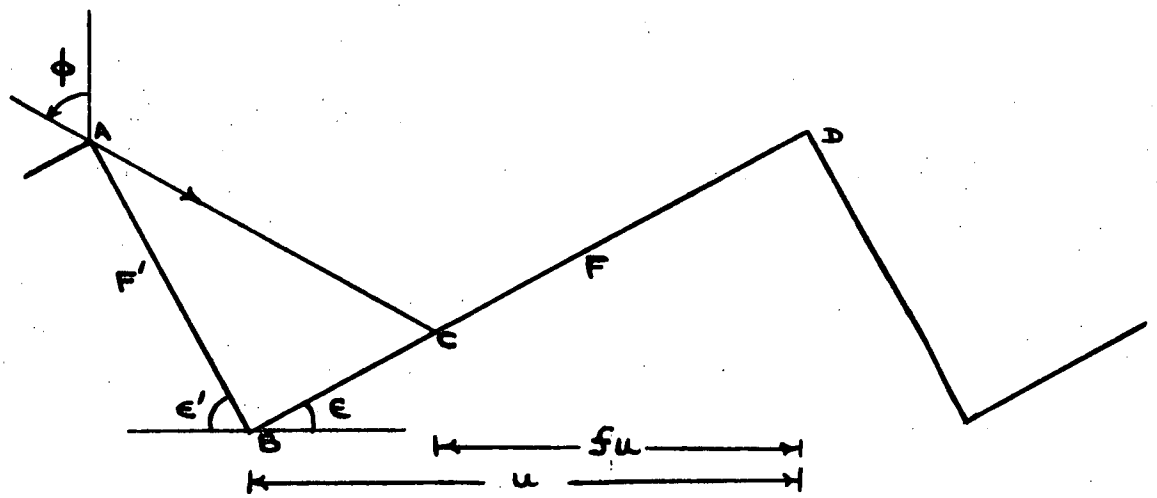


Figure (2.3). The shadowing of part of the triangular profile of a diffraction grating is illustrated, for the case when $\phi > 90^{o}-\varepsilon'$. The section ABC of the groove surface is in shadow, and is not included in the calculation of the diffracted field.

The calculation of the foreshortening factor f is made by assuming that the incident and diffracted rays travel along rectilinear paths within the grating grooves - a region whose dimensions are typically of the order of the wavelength of the radiation. This assumption has led to the description by G. Cerutti-Maori and R. Petit (2.4) of this formulism as "a curious mixture of geometrical optics and of physical optics".

In the evaluation of the diffraction efficiency, only the first scattering of the incident wave by the grating was taken into account. The contribution to the final diffracted field from radiation diffracted from one groove facet to another facet and there re-diffracted was neglected. Such secondary diffraction effects have been taken into account by Janot and Hadni (2.5). Palmer and Le Brun (2.6) have included secondary and tertiary diffraction effects in their calculation.

If higher order diffraction effects are significant, then an accurate evaluation of the final diffracted field must take into account contributions from all orders of diffraction, expressing the efficiency as the sum of an infinite series whose terms describe primary, secondary, tertiary, quaternary,... effects. No current scalar theory follows such a procedure.

The above scalar, single-scattering formulae for diffraction efficiency were derived for triangular profile gratings. Corresponding expressions have been given by J.L. Uretsky (2.7) and by J.A. Aas (2.8) for gratings having a sinusoidal profile.

Consider the formula (2.5) giving the diffraction efficiency for the triangular profile in the absence of shadowing. The efficiency function is oscillatory in nature, with an envelope given by

$$B = \left[\frac{1 + x^2}{(1-cx)(1+c'x)} \right]^2 \frac{\lambda^2 (c+c')^2}{4\pi^2 a^2 \cos\phi \cos\psi} \quad . \qquad (2.12)$$

The envelope has asymptotes for those angles of incidence such that

$$x = \cot\left(\frac{\phi+\psi}{2} \right) = \frac{1}{c} = \cot(\varepsilon) \qquad (2.13)$$

or

$$x = \cot\left(\frac{\phi+\psi}{2} \right) = -\frac{1}{c'} = \cot(-\varepsilon'). \qquad (2.14)$$

These asymptotes correspond to angles of incidence such that, in the order considered, the diffracted ray is obtained from the incident ray by specular reflection from either groove facet. When the grating is used in this configuration, it is said to be "on-blaze". The scalar theory indicates that a grating is most efficient when used near the blaze position. For S polarized radiation, rigorous electromagnetic theory (2.9, 2.10) confirms this result for gratings with rectangular grooves.

A computer program has been written to calculate scalar values of diffraction efficiencies for the triangular profile, based on equations (2.5) to (2.11). The efficiencies are calculated for all propagating orders, which for a given wavelength (λ) and angle of incidence (ϕ) are comprised between the limits

$$LL = -\text{ entier } \left(\frac{a}{\lambda} (1 - \sin\phi) \right) \qquad (2.15)$$

and

$$UL = \text{entier } \left(\frac{a}{\lambda} (1 + \sin\phi) \right) \quad . \qquad (2.16)$$

The total diffracted energy is also calculated, as a means of assessing the accuracy of the computed efficiency values:

$$D.E. = \sum_{m=LL}^{UL} E(m) \quad . \qquad (2.17)$$

Since the energy of the incident wave is normalized to unity, conservation of energy dictates that D.E. should also be equal to one. The greater the deviation of the diffracted energy is from unity, the greater is the inaccuracy of the scalar efficiency values.

The program was used to investigate the circumstances under which the scalar theory can be expected to give accurate results. The only previous systematic study of this question was made by G. Cerutti-Maori and R. Petit (2.4). For a grating with facet angles 20^{0} and 70^{0}, they found that the scalar theory was inaccurate even for relative wavelength (λ/d) values smaller than 0.1. The theory was most accurate near the blaze position of the grating. Comparison of theory with experiment showed qualitative agreement to exist for a grating with a blaze angle of 10^{0} (i.e., the smaller groove facet angle was 10^{0}). For a blaze angle of $26^{0}30'$, not even qualitative agreement was obtained.

Figures (2.4) to (2.7) show the variation of the diffracted energy values furnished by the scalar theory program with normalized wavelength. The figures correspond to gratings having rectangular grooves and blaze angles of 2.5^{0}, 5.0^{0}, 10.0^{0} and 20.0^{0}. For each grating, curves associated with constant angles of incidence (ϕ) of 20^{0}, 0^{0} and -20^{0} are given.

Figure (2.4) shows the behaviour of the theory for a shallow profile with groove angles 2.5^{0} and 87.5^{0}. The scalar calculations are of reasonable accuracy for this grating, with the diffracted energy curve for $\phi = -20^{0}$ having the largest range of variation (from 0.94 to 1.15). The most prominent feature of the curves is their segmentation by Wood anomalies. The curves tend to slope upwards as λ tends to λ_R, and then fall rapidly at λ_R, before starting to rise again.

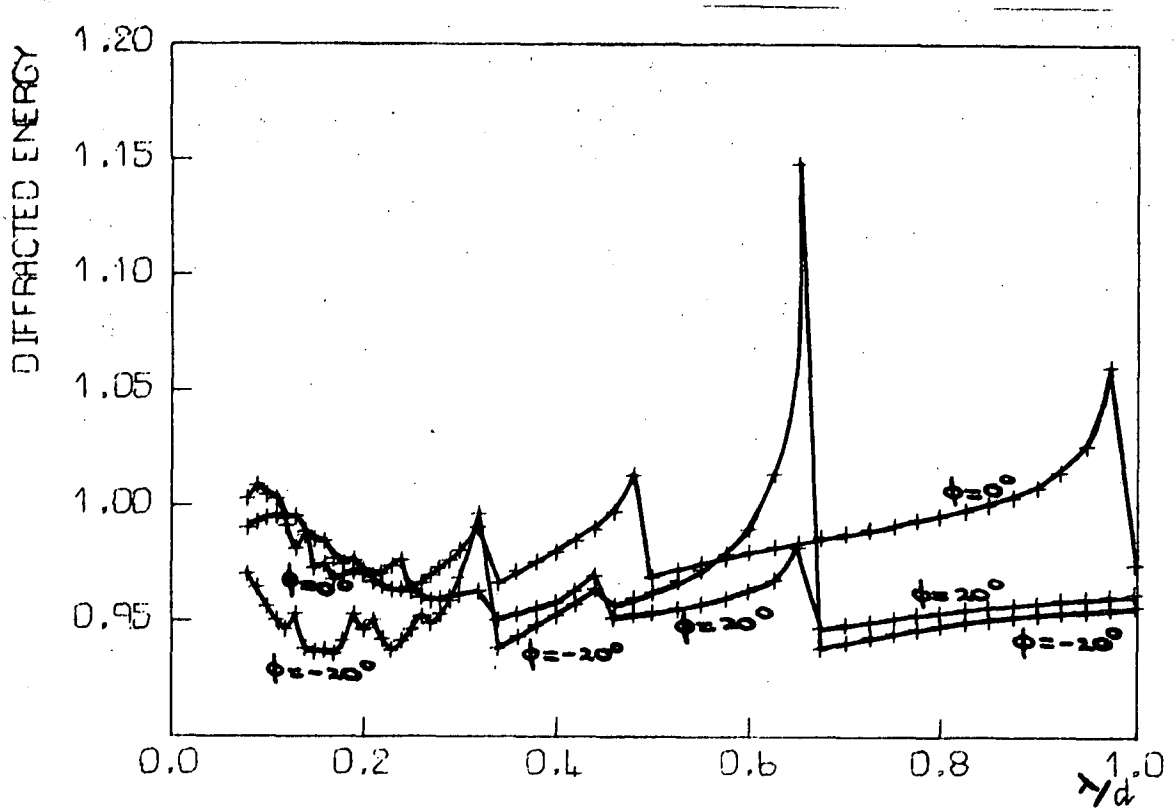


Figure (2.4). The three curves correspond to a triangular profile grating having facet angles 2.5^{o} and 87.5^{o}, and show the variation of diffracted energy with normalized wavelength, for three values (-20^{o}, 0^{o} and 20^{o}) of the angle of incidence ϕ.

The largest discontinuity occurs near $\lambda/d = 0.65$, for $\phi = -20^{o}$. The efficiency in the order +1 rises to over 20%, just before the order passes off. Since the scalar calculation does not couple the efficiency of the grazing order to the strengths of the other orders, the diffracted energy curve rises up rapidly near λ_R, and then falls abruptly when the order passes off and its efficiency goes to zero. Evidently, if the criterion of conservation of energy is to be always satisfied by the results of a theory, that theory should couple together the efficiencies of all orders through an appropriate boundary condition, so that a rapid rise in the strength of one order can be counterbalanced by a fall in the strength of another. The scalar theory is not based on the satisfaction of a rigorous boundary condition, and so its results give a diffracted energy value which undergoes local fluctuations as the efficiencies of the various orders fluctuate.

Figure (2.5) corresponds to a profile with groove angles 5^{o} and 85^{o}. With this deeper profile, the diffracted energy curves undergo variations which are larger than in the previous case. The segmented nature of the curves due to the passing off of spectral orders is again clear.

For a blaze angle of 10^{o} (Figure (2.6)), the scalar theory is of inadequate accuracy for angles of incidence of 0^{o} and -20^{o}. For $\phi = 20^{o}$, the diffracted energy varies between 0.84 and 1.02, so that the calculated efficiency values for this case might give a qualitative indication of the behaviour of the grating. For $\phi = 0^{o}$ and $\phi = -20^{o}$, the limits of variation are respectively 0.67 to 1.14, and 0.48 to 0.94. For these two angles of incidence, scalar efficiency values would not have even qualitative significance.

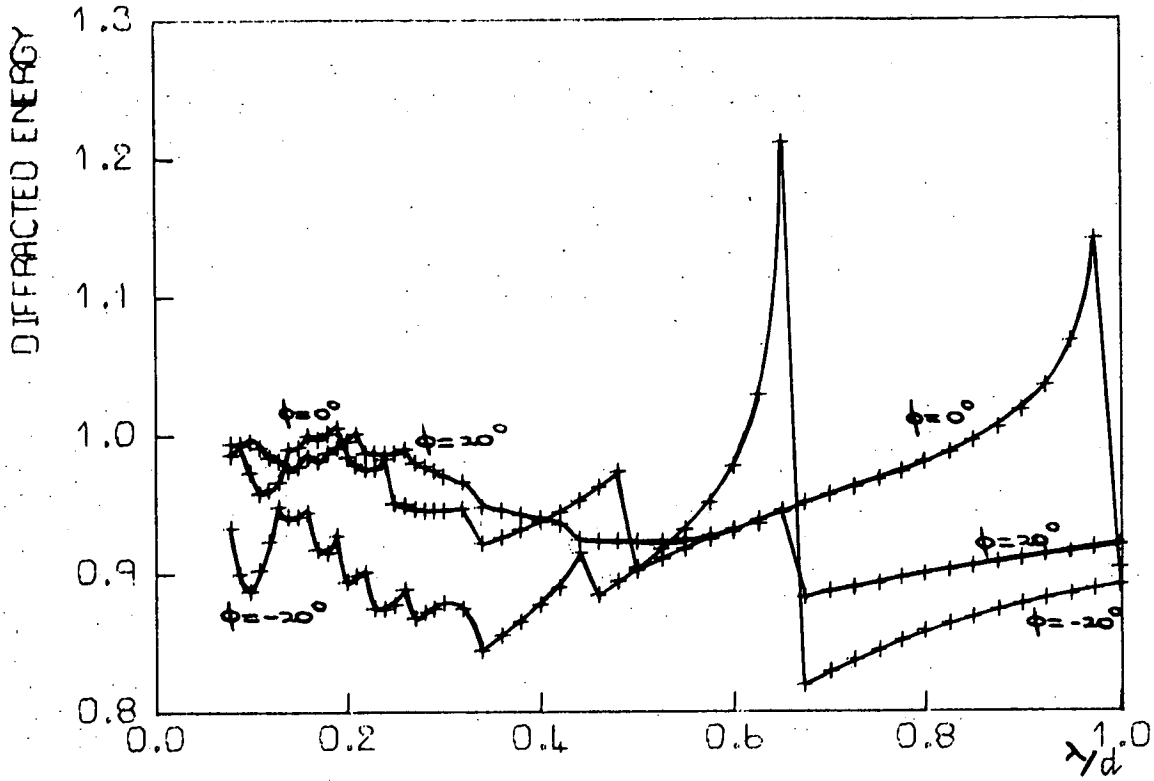


Figure (2.5). The curves correspond to those of figure (2.4), except that the grating now has facet angles 5.0^{o} and 85.0^{o}.

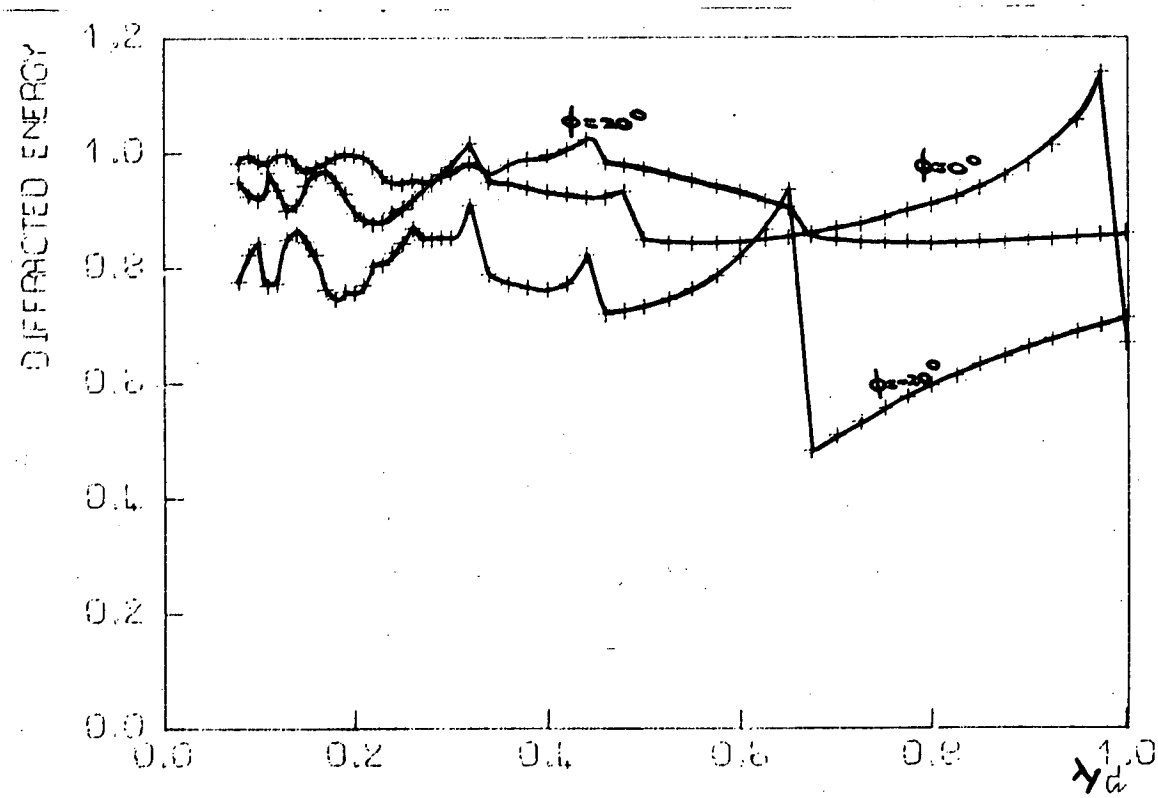


Figure (2.6). The curves correspond to those of figure (2.4), except that the grating now has facet angles 10.0^{o} and 80.0^{o}.

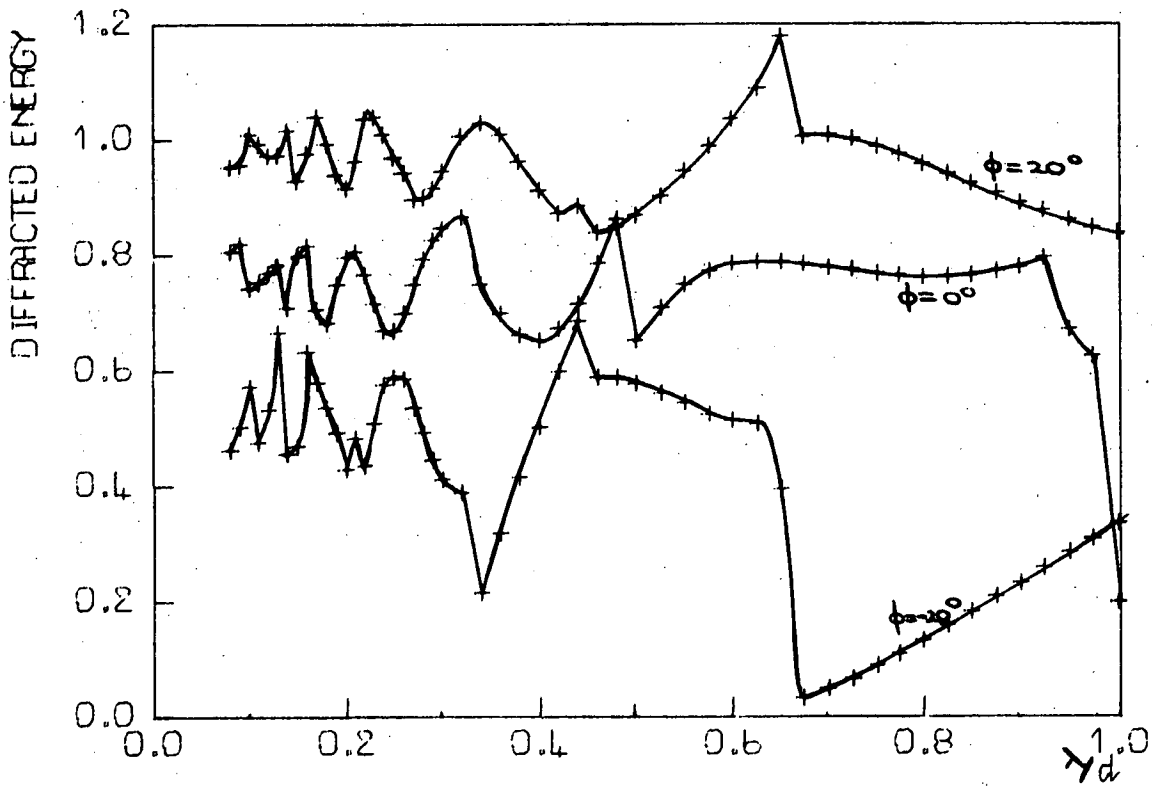


Figure (2.7). The curves correspond to those of figure (2.4), except that the grating now has facet angles 20.0^0 and 70.0^0.

The curves for a blaze angle of 20^{o} (Figure (2.7)) correspond to those investigated by Cerutti-Maori and Petit (2.4). The scalar theory obviously does not become accurate when the wavelength becomes small compared with the grating period. Only the curve for $\phi = 20^{o}$ moves towards unity as λ/d decreases towards 0.1. It varies overall between the limits 0.84 and 1.18. The diffracted energy value for $\phi = -20^{o}$ oscillates around 0.5 as λ/d tends to 0.1.

In summary, the formulation of Madden and Strong (2.1) is of reasonable accuracy when applied to triangular profile gratings only if the blaze angle does not exceed 5^{o}. The diffracted energy values tend to fluctuate rapidly in the vicinity of Rayleigh wavelengths. For other than shallow profiles, they do not tend towards unity as λ/d decreases. Since the polarization of diffracted light by gratings decreases as the period becomes coarse when compared with the wavelength (2.11), a scalar theory should be able to describe adequately the diffraction process for small values of λ/d. Such a theory should afford diffracted energy values close to unity for coarse rulings. It is evident that the formulation we have described does not satisfy this requirement, and is thus not the scalar theory desired for dealing with diffraction by coarse gratings (2.4).

2.3. A METHOD BASED ON THE RAYLEIGH EXPANSION

The representative of the group of theories of grating diffraction based on the Rayleigh expansion which we will consider in this section is the method of R. Petit (2.12-2.14). The formulism will be given for the case when the incident field is P polarized. The calculation of S polarization efficiencies follows along similar lines.

Consider a P polarized wave incident at an angle θ on a perfectly conducting grating of period d, whose profile is described by a function f(x). The Rayleigh expansion for the z component of the diffracted electric field strength is

$$E^d = \sum_{n=-\infty}^{\infty} B_n \exp[ik(x \sin \theta_n + y \cos \theta_n)] = \sum_{n=-\infty}^{\infty} B_n E_n^d . \quad (2.18)$$

Here spatial coordinates are specified with respect to a system Oxyz, where the Oz axis is parallel to the grating grooves and the Ox axis lies in the plane of the grating blank, perpendicular to the line of the grooves.

The angle of diffraction of the nth order is θ_n, where

$$\sin \theta_n = \frac{n\lambda}{d} + \sin \theta , \quad (2.19)$$

in accordance with the grating equation. Since the diffracted field must take the form of a sum of progressive waves and evanescent waves which become attenuated as they move away from the grating,

$$\cos \theta_n = (1 - \sin^2\theta_n)^{\frac{1}{2}} \text{ if } |\sin \theta_n| \leq 1; \quad (2.20)$$

and

$$\cos \theta_n = i(\sin^2\theta_n - 1)^{\frac{1}{2}} \text{ if } |\sin \theta_n| > 1. \quad (2.21)$$

If E^i denotes the z component of the electric field strength of the incident wave, then the boundary condition on the grating surface is

$$E^d[x, f(x)] = -E^i[x, f(x)] = g(x) . \quad (2.22)$$

Using (2.18), this boundary condition becomes

$$g(x) = \sum_{n=-\infty}^{\infty} B_n E_n[x, f(x)] = \sum_{n=-\infty}^{\infty} B_n \psi_n . \quad (2.23)$$

The unknown quantities it is desired to evaluate are the amplitudes B_n of the various orders. The functions $g(x)$ and ψ_n are known. The problem is thus reduced to one of finding the coefficients in the expansion of a known function in terms of an infinite set of known functions. Petit has found that the latter problem can best be solved using the Gram-Schmidt orthogonalization process (2.15).

In the space of piecewise-continuous functions of period d, the inner product of two functions f_1 and f_2 is defined as

$$(f_1, f_2) = \int_0^d f_1(x)\, f_2^*(x)\, dx \ . \qquad (2.24)$$

The norm of a function f is

$$\|f\| = \int_0^d |f(x)|^2 dx \ . \qquad (2.25)$$

By changing the method of counting of diffracted orders, we will restrict the right-hand side of equation (2.23) to be a sum over positive integers. From the ψ_n, we use the Gram-Schmidt process to construct an orthogonal set $\phi_1, \phi_2, \ldots, \phi_n, \ldots$. Then

$$\phi_r = \psi_r - \sum_{i=1}^{r-1} \frac{(\psi_r, \psi_i)}{\|\phi_i\|}\, \phi_i = \psi_r + \sum_{i=1}^{r-1} \lambda_r^i\, \phi_i \ . \qquad (2.26)$$

An orthonormal set of functions $x_1, x_2, \ldots, x_n, \ldots$ can also be constructed, where

$$x_i = \phi_i / \sqrt{\|\phi_i\|} \ . \qquad (2.27)$$

We now expand $g(x)$ in terms of the x_i :

$$g(x) = \sum_{i=1}^{\infty} (g, x_i) x_i = \sum_{i=1}^{\infty} c_i\, x_i \ . \qquad (2.28)$$

Since the diffraction problem is to be solved numerically, the infinite

summations must be truncated - say at an index p. Let

$$g_p(x) = \sum_{i=1}^{p} c_i \, x_i \, . \qquad (2.29)$$

To determine a suitable value of p, an accuracy tolerance δ is set on the expansion of the function g. The construction process of the x_i is stopped when this tolerance is achieved - i.e., when

$$\delta^2 > \|g-g_p\| = \|g\| - \sum_{i=1}^{p} c_i \, c_i^* \, . \qquad (2.30)$$

The two expansions (2.23) and (2.29) of g are then compared. This (2.13) leads to the following set of equations

$$c_q / \sqrt{\|\phi_q\|} = B_q - \sum_{n=q+1}^{p} B_n \, \lambda_n^q \, . \qquad (2.31)$$

This set of equationscan be solved explicitly:

$$B_p = c_p / \sqrt{\|\phi_p\|} \, , \; B_{p-1} = B_p \, \lambda_p^{p-1} + c_{p-1} / \sqrt{\|\phi_{p-1}\|} \, , \ldots \qquad (2.32)$$

Equation (2.32) enables the evaluation of all the unknown amplitudes of the components of the diffracted field (up to the index p).

The numerical execution of this method of solution requires the finding of inner products of the form (ψ_i, ψ_j) and (g, ψ_j). These can all be expressed in terms of the following integral:

$$I(t,u,v) = \int_0^d \exp(-ku \, f(x)) \exp[i(tKx + k \, f(x)v)], \qquad (2.33)$$

where $K = 2\pi/d$, t is integral and u and v are reals ($u \geq 0$). This single integral can be evaluated using a quadrature rule. In a general profile program which has been written to provide Rayleigh expansion efficiency values for an arbitrary function f(x), it has

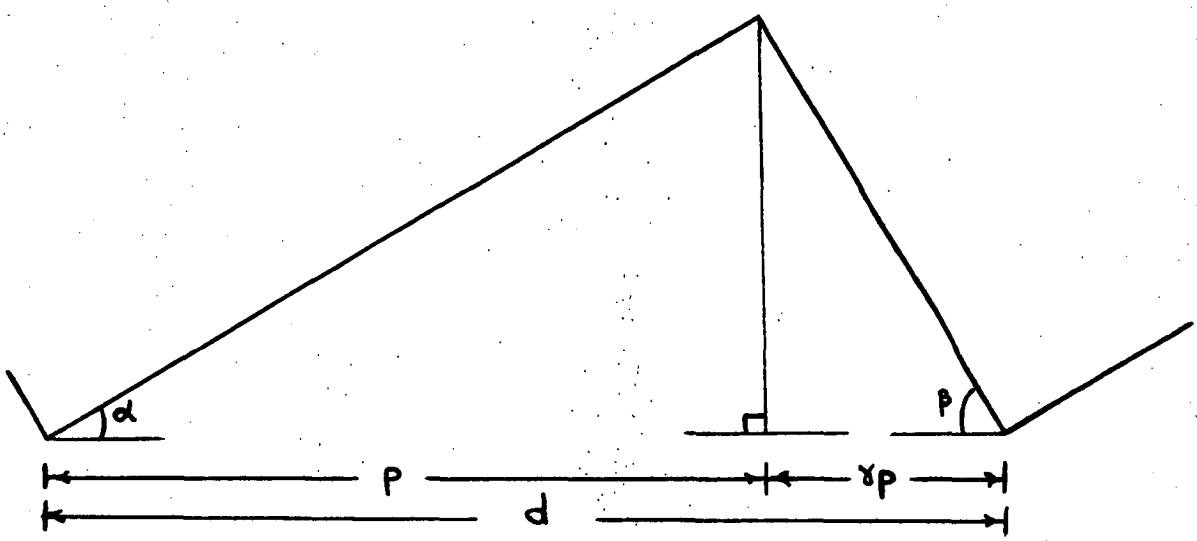


Figure (2.8). Definitions of quantities characterizing the profile of a grating having triangular grooves. These are the period (d), the facet angles (α, β), the projection of the longer groove facet onto the plane of the blank (p) and $\gamma = (d-p)/p$.

been found convenient to use Filon's rule (2.16).

Analytical expressions for $I(t,u,v)$ can be derived for gratings having sinusoidal and triangular profiles. In the case of the triangular profile with facet angles α and β shown in Figure (2.8),

$$I(t,u,v) = \frac{[1-\exp(-kup \tan\alpha) \exp ip(Kt+kv \tan\alpha)]}{[ku \tan\alpha - i(Kt+kv \tan\alpha)]}$$

$$+ \frac{\gamma[1-\exp(-kup \tan\alpha) \exp ip(-\gamma Kt+kv \tan\alpha)]}{[ku \tan\alpha - i(-\gamma Kt+kv \tan\alpha)]} . \quad (2.34)$$

For a sinusoidal profile with equation

$$f(x) = A(1 - \cos Kx) , \quad (2.35)$$

it can be shown that

$$I(t,u,v) = \frac{2\pi}{K} \exp i[kA(v+iu) - \frac{t\pi}{2}] J_t[kA(v+iu)], \quad (2.36)$$

where J_t denotes the Bessel function of the first kind, of order t.

Programs utilizing the expressions (2.34) and (2.36) have been written to calculate diffraction efficiencies for triangular and sinusoidal profile gratings. In these programs, the construction of the orthogonal functions is limited by a parameter L - if the accuracy tolerance has not been achieved by p = L, the rest of the calculation proceeds with p taking this value. Values of L up to 30 are used. The typical value for the accuracy tolerance is 0.001. The triangular profile program calculates both P and S polarization efficiencies. For the sinusoidal profile, separate programs are used for the two polarizations.

The triangular profile program was used to calculate diffracted energy curves for P and S polarizations for a profile with facet angles 10^0 and 80^0. These curves are shown in Figure (2.9), and

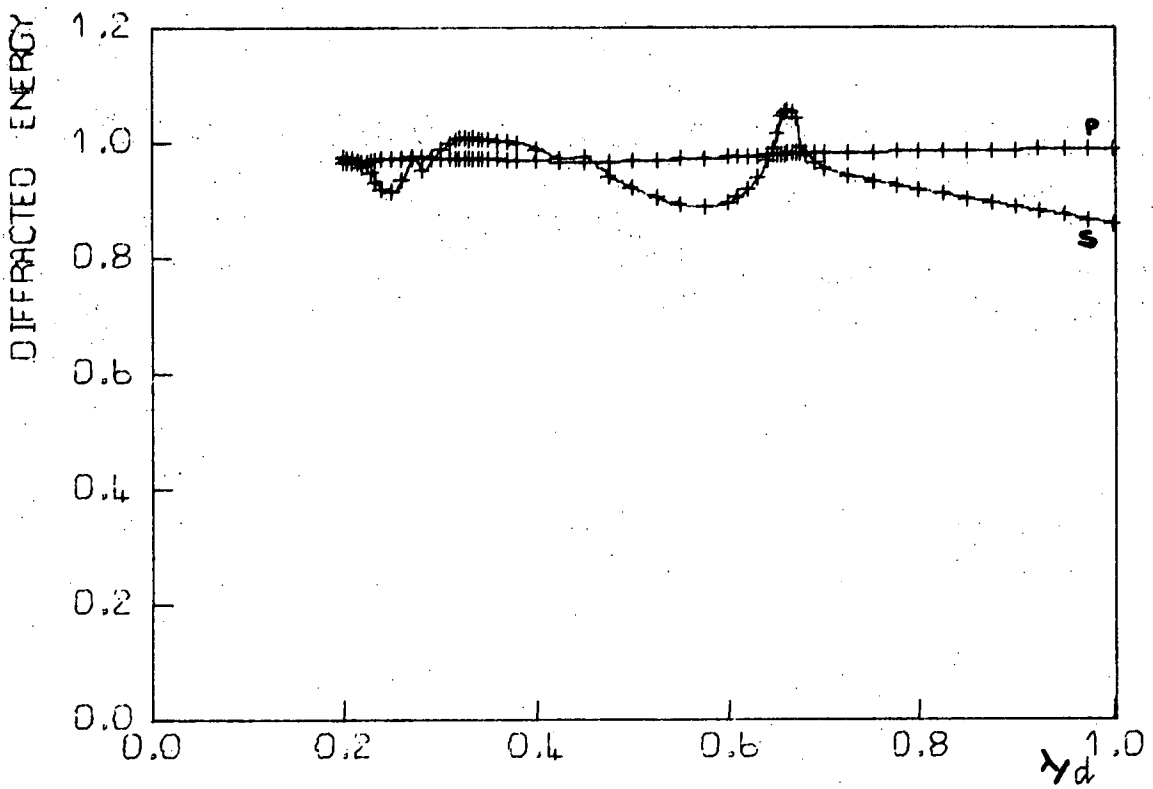


Figure (2.9). Diffracted energy curves computed using a Rayleigh theory formulation, for a grating having a triangular profile with facet angles 10^0 and 80^0, used at a constant angle of incidence ($\theta = 20^0$). Curves for both P and S polarized radiation are shown.

should be compared with the corresponding scalar curve (that for $\phi = -20^0$ of Figure (2.6)). The scalar theory gave diffracted energy values varying between 0.48 and 0.94. For the Rayleigh theory and P polarization, the diffracted energy varies smoothly with wavelength between the limits of 96% and 99%. The S polarization diffracted energy behaves in a more oscillatory fashion with wavelength, but always lies between the limits of 86% and 106%.

The Rayleigh theory couples all the intensities of diffracted orders, in order to satisfy the appropriate boundary condition on the grating surface. Thus, the passing off of one order affects the efficiencies of other orders. The theory can (and does) give Wood anomalies in the propagating orders, and so can preserve a reasonable energy balance in the wavelength regions in which strong grazing orders pass off (in contrastwith the scalar theory).

The smoothness of the P polarization diffracted energy curve is a consequence of the weak nature of P Wood anomalies and grazing orders for shallow profiles.

The results of the second test of the Rayleigh theory for triangular profiles are shown in Figure (2.10). The diffracted energy curves correspond to gratings with symmetrical profiles ($\alpha = \beta$), used with a normally incident beam of wavelength $\lambda/d = 0.4368$. The theoretical efficiency values obey well the criterion of conservation of energy until α reaches 30^0. There, for P polarization, the diffracted energy value jumps to 1.53. The corresponding S polarization loss of accuracy occurs at $\alpha = 34^0$.

The superiority of the Rayleigh theory over the scalar formulation is evident from Figures (2.9) and (2.10). The former is useful for profiles with blaze angles up to 30^0, whereas the latter was

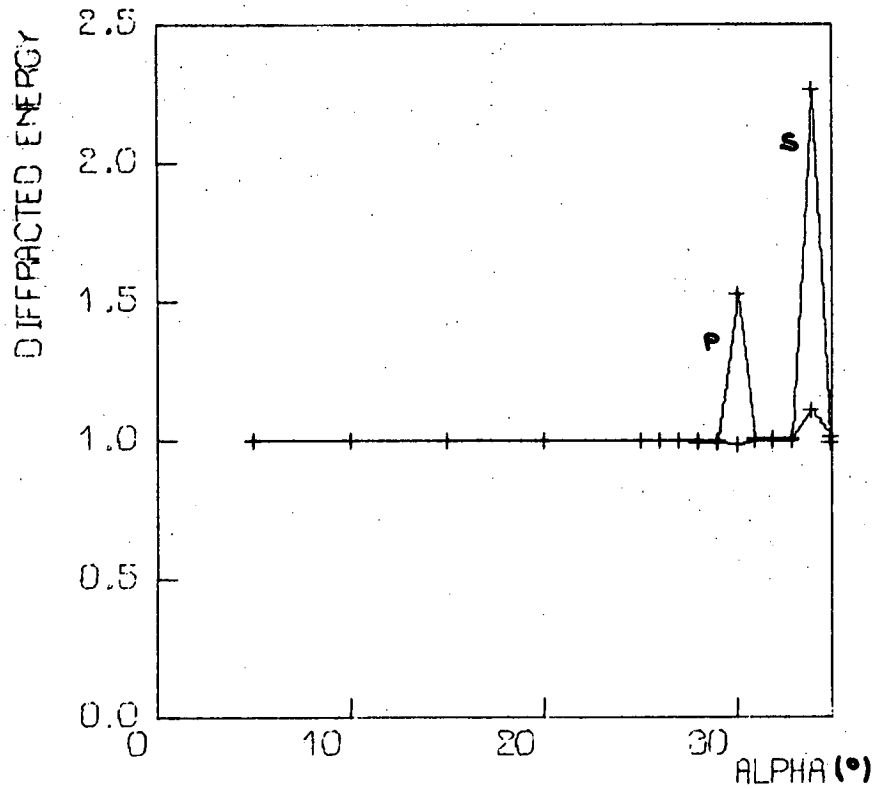


Figure (2.10). The variation of diffracted energy as a function of the blaze angle (α) in degrees is shown, for gratings having symmetrical triangular profiles, used with normally incident light of wavelength $\lambda/d = 0.4368$. (Rayleigh theory results.)

limited to blaze angles up to 5^{o}.

Diffracted energy curves for the sinusoidal profile specified by equation (2.35) are shown in Figures (2.11) and (2.12). The first of these corresponds to gratings used with normally incident light of wavelength $\lambda/d = 0.4368$. The P polarization curve is similar to one given by R. Petit and M. Cadilhac (2.17). The value of L used in the calculations was fifteen. For $A/d > 0.16$, the increase of L to thirty did not improve the accuracy of the results. As noted by Petit and Cadilhac, the value of A/d can exceed the theoretical upper limit of validity of the Rayleigh expansion (which is $u_0/2\pi = 0.072$) without the results losing their physical usefulness. The accuracy of the S polarization calculations is in general worse than that for P polarization. The respective values of A/d for which the P and S polarization diffracted energy values decrease below 0.95 are 0.16 and 0.14.

The theoretical analysis of the region of validity of the Rayleigh theory indicates a limit which is independent of wavelength. The above conclusions refer to $\lambda/d = 0.4368$. Consider the situation for a wavelength of 0.21 (Figure (2.12)).

For this smaller wavelength, with more orders being formed, the diffracted energy decreases more rapidly with increasing groove amplitude. The results for P and S polarization obtained with L = 15 indicate that serious energy imbalances begin to occur after $A/d = 0.10$. For P polarization the diffracted energies for $A/d = 0.10$ and 0.12 are respectively 99% and 93%. For S polarization, the corresponding values are 99% and 94%. Thus, with the decrease of the wavelength, the limit of good accuracy of the program has neared the theoretical limit of validity derived by Petit and Cadilhac (2.17).

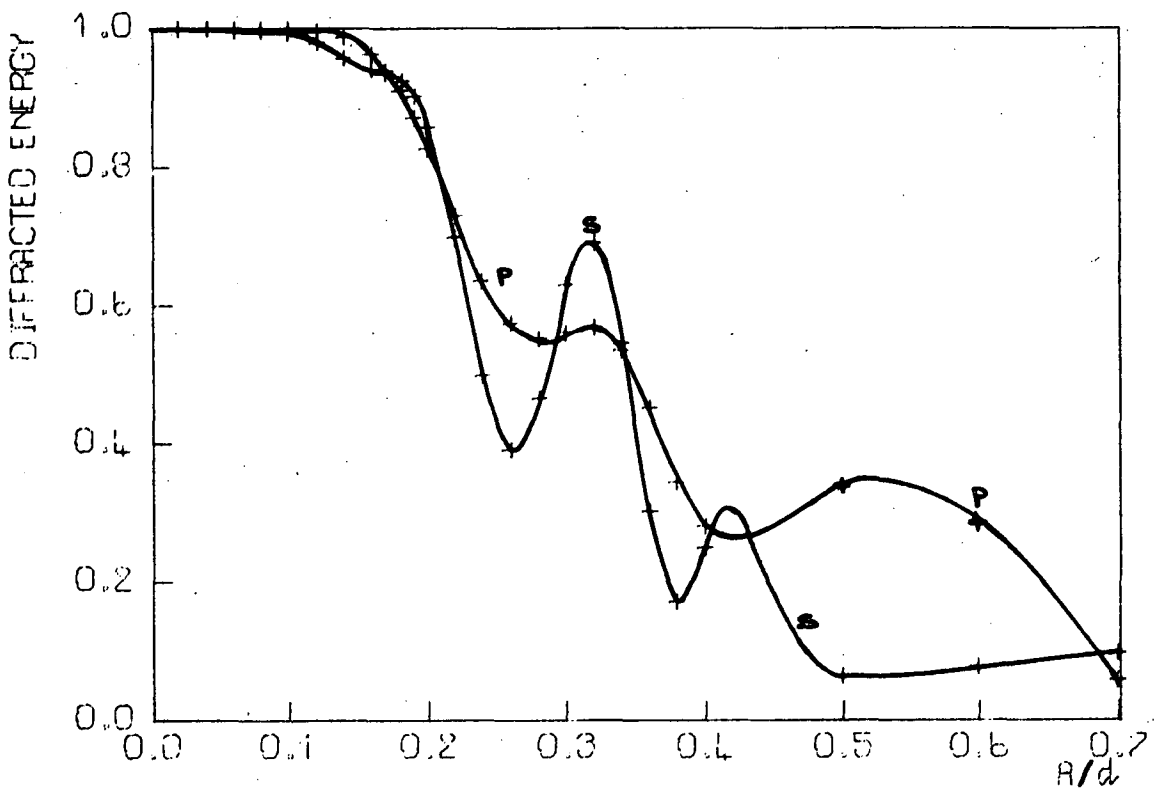


Figure (2.11). The variation of diffracted energy as a function of normalized groove amplitude is shown, for sinusoidal gratings used with normally incident light of wavelength $\lambda/d = 0.4368$. (Rayleigh theory results.)

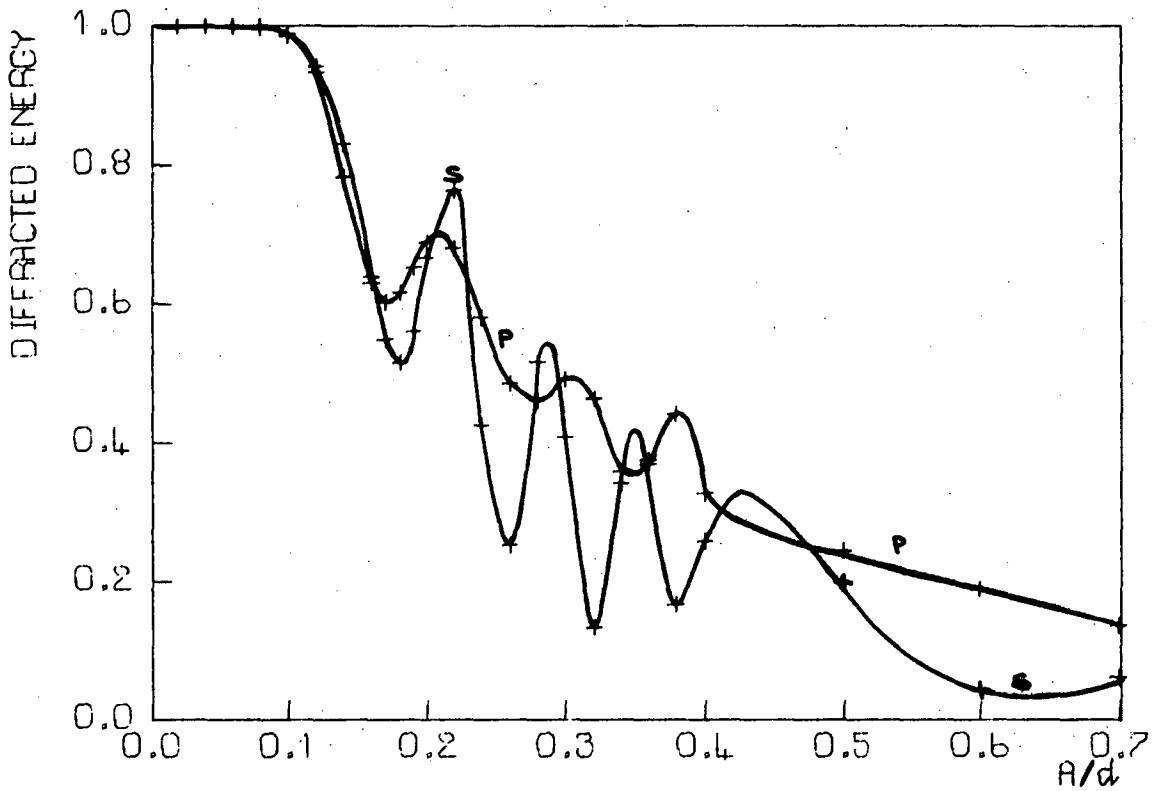


Figure (2.12). The variation of diffracted energy as a function of normalized groove amplitude is shown, for sinusoidal gratings used with normally incident light of wavelength $\lambda/d = 0.21$. (Rayleigh theory results.)

2.4. FOURIER SERIES METHODS

In this section, the various methods developed by R. Petit in association with M. Cadilhac based on the solution of integral equations using Fourier series techniques will be discussed. These were the first rigorous methods devised for the numerical solution of the two-dimensional problem of the diffraction by a perfectly conducting grating.

Consider first the formulation for the case of a P polarized incident wave, which was developed in references (2.18-2.20). The notation used is that of the previous section, with a plane wave incident at an angle θ on a grating whose surface is specified by the equation $y = f(x)$. All fields are specified in terms of their components along the Oz axis. The incident field has a spatially dependent part

$$E^i = \exp[ik(x \sin \theta - y \cos \theta)] \qquad (2.37)$$

and a time-dependent part $\exp(-i\omega t)$. Since the time dependence of the incident and diffracted fields is the same, it can be omitted in the development of the formulism.

The diffracted field $E^d(x,y)$ is defined for $y \geq f(x)$ and has the properties

$$\Delta E^d + k^2 E^d = 0 , \qquad (2.38)$$

and

$$E^d(x, f(x)) = -E^i(x, f(x)) = g(x) . \qquad (2.39)$$

As well, E^d must behave as a sum of outgoing waves and remain finite as $y \to +\infty$. From the periodic properties of the grating, and using equation (2.37),

$$E^d(x+d, y) = E^d(x, y) \exp(ikd \sin \theta) . \qquad (2.40)$$

For convenience, the function E^d is prolonged into the region $y < f(x)$ by a function $E^p(x,y)$. Then E^p will satisfy equations (2.38), (2.39) and (2.40), and behave as a sum of outgoing waves and remain finite as $y \to -\infty$. It can easily be seen that $-E^i(x,y)$ is the unique choice for the prolonging function E^p.

We can now consider the function $\tilde{E}(x,y)$ equal to E^d if $y \geq f(x)$ and to E^p if $y < f(x)$. By definition, $\tilde{E}$ is everywhere continuous. However, nothing guarantees the continuity of its derivatives across the grating surface. Petit has shown (2.20) that if $\tilde{E}$ is regarded as a distribution, it satisfies the modified Helmholtz equation

$$\Delta\tilde{E} + k^2 \tilde{E} = 2ik\, \phi(x)\, \delta(y - f(x)) \ . \qquad (2.41)$$

The unknown function $\phi(x)$ is proportional to the current density induced on the grating surface by the incident field.

From equation (2.40) and its analogue for E^p, it can be shown that

$$\tilde{E}(x+d,y) = \tilde{E}(x,y) \exp(ikd \sin\theta) \ , \qquad (2.42)$$

and

$$\phi(x+d) = \phi(x) \exp(ikd \sin\theta) \ . \qquad (2.43)$$

Hence, these two functions have Fourier series of the form

$$\tilde{E}(x,y) = \sum_{n=-\infty}^{\infty} \tilde{E}_n(y)\, u_n(x) \ , \qquad (2.44)$$

and

$$\phi(x) = \sum_{n=-\infty}^{\infty} \phi_n\, u_n(x) \ , \qquad (2.45)$$

where the u_n are harmonic functions:

$$u_n = \exp[i(nK + k \sin\theta)x] \ . \qquad (2.46)$$

The use of these Fourier series in the modified Helmholtz equation (2.41) leads to the following differential equation:

$$\frac{d^2\tilde{E}_n}{dy^2} + \Omega_n \tilde{E}_n = \frac{ikK}{\pi} \int_0^d \phi(x)\, \delta(y-f(x))\exp[-i(nK+k \sin\theta)x]dx \ , \tag{2.47}$$

where

$$\Omega_n = k^2 - (nK + k \sin\theta)^2 \ . \tag{2.48}$$

If U denotes the set of integers such that Ω_n is positive, and if

$$\begin{aligned} x_n &= \sqrt{\Omega_n} \quad \text{for } n \in U \\ &= i\sqrt{-\Omega_n} \quad \text{for } n \notin U \ , \end{aligned} \tag{2.49}$$

then the solution of equation (2.47) is

$$\tilde{E}_n(y) = \frac{kK}{2\pi x_n} \int_0^d \phi(x') \exp[-i(nK+k \sin\theta)x'+ix_n|y-f(x')|]dx'. \tag{2.50}$$

Hence, we have

$$\tilde{E}(x,y) = \sum_{n=-\infty}^{\infty} \frac{kK}{2\pi x_n} \int_0^d \phi(x')\exp[i(nK+k \sin\theta)(x-x')+ix_n|y-f(x')|]dx', \tag{2.51}$$

or, in more condensed form,

$$\tilde{E}(x,y) = \int_0^d R(x-x',\ y - f(x'))\ \phi(x')\ dx' \ , \tag{2.52}$$

where

$$R(x,y) = \frac{kK}{2\pi} \sum_{n=-\infty}^{\infty} \frac{\exp[i(nK+k \sin\theta)x+ix_n|y|]}{x_n} \ . \tag{2.53}$$

Since the boundary condition on the grating surface is

$$\tilde{E}(x,\ f(x)) = g(x) \ , \tag{2.54}$$

equation (2.53) leads to the following Fredholm equation:

$$\int_0^d \rho(x,x')\ \phi(x')\ dx' = g(x)\ . \qquad (2.55)$$

Here

$$\rho(x,x') = R(x-x',\ f(x)-f(x'))\ . \qquad (2.56)$$

The kernel function ρ has singularities at points (x,x') such that $f(x) = f(x')$. However, it has been shown (2.20) that ρ is integrable over the interval $[0,d]$.

The Fredholm equation (2.55) is soluble using a Fourier series technique. Using a truncated series having $P = 2Q+1$ terms, we express $g(x)$ as:

$$g(x) \simeq \frac{1}{d} \sum_{j=-Q}^{Q} g_j\ u_j(x)\ , \qquad (2.57)$$

where

$$g_j = -\int_0^d \exp[-ij\ Kx - ik\ f(x)\ \cos\ \theta]dx\ . \qquad (2.58)$$

The unknown function ϕ is expressed in terms of a similar series:

$$\phi(x) \simeq \sum_{j=-Q}^{Q} \phi_j\ u_j(x)\ . \qquad (2.59)$$

An approximate expression is also used for R, having $S = 2M+1$ terms:

$$R(x,y) \simeq \frac{kK}{2\pi} \sum_{n=-M}^{M} \frac{\exp[i(nK + k\ \sin\ \theta)x + ix_n|y|]}{x_n}\ . \qquad (2.60)$$

The kernel function ρ is a function of two variables, and thus is expressed as a Fourier series over x (with running index j) and also over x' (with running index j'). The typical coefficient is

$$\rho_{jj'} = \int_0^d\int \rho(x,x')\ \exp[-i(jK+k\ \sin\ \theta)x+i(j'K+k\sin\ \theta)x']dx\ dx'. \qquad (2.61)$$

Using equation (2.60),

$$\rho_{jj'} = \sum_{n=-M}^{M} \rho_{jj'n} \qquad (2.62)$$

where

$$\rho_{jj'n} = \frac{kK}{2\pi} \int_0^d\int \frac{\exp(-\alpha_{nj}x+\alpha_{nj'}x'+ix_n|f(x)-f(x')|)}{x_n} \, dx \, dx' \quad (2.63)$$

if $$\alpha_{nj} = i(j-n)K \quad , \quad \alpha_{nj'} = i(j'-n)K \; . \quad (2.64)$$

Substitution of the Fourier series for ρ, ϕ and g into the integral equation (2.55) and use of the orthogonality properties of harmonic functions results in the following set of linear equations:

$$\sum_{j'=-Q}^{Q} \phi_{j'} \rho_{jj'} = g_j \; , \quad (j = -Q,\ldots,0,\ldots+Q). \quad (2.65)$$

Once this set of equations is solved for the unknown coefficients $\phi_{j'}$, the amplitude of the nth order component of the diffracted field can be found using the expression

$$B_n = \frac{kK}{2\pi x_n} \sum_{j'=-Q}^{Q} \phi_{j'} \int_0^d \exp[\alpha_{nj'}x' - ix_n f(x')]dx' \; . \quad (2.66)$$

As a check on the calculations, the nth order component of E^p can also be found:

$$C_n = \frac{kK}{2\pi x_n} \sum_{j'=-Q}^{Q} \phi_{j'} \int_0^d \exp[\alpha_{nj'}x'+ix_n f(x')]dx' . \quad (2.67)$$

The calculated values of the C_n should be such that

$$\begin{aligned} C_0 &\simeq 1 \\ C_n &\simeq 0 \text{ if } n \neq 0 \; . \end{aligned} \quad (2.68)$$

The numerical evaluation of efficiencies requires the calculation of the g_j (from the single integral of equation (2.58)), the calculation of the $\rho_{jj'n}$ (from the double integral of equation (2.63)), the inversion of the system of linear equations (2.65) and the final

evaluation of the B_n and C_n from the single integrals (2.66) and (2.67). While in principle these steps can all be performed numerically for an arbitrary profile function f(x), the calculation of the $\rho_{jj'n}$ involves a prohibitive quantity of computer time unless an analytic expression can be used to evaluate this double integral of an oscillatory function. This requirement restricts the application of the Fourier series method to profiles composed of straight line segments. Analytic expressions for the g_j and $\rho_{jj'n}$ in the case of the triangular profile of Figure (2.8) have been given by R. Petit (2.20).

The development of a formulism for the second fundamental polarization of the incident wave (that in which the magnetic field vector is parallel to the grating grooves) involved greater difficulty than the construction of the above P polarization solution. A method suggested by R. Petit (2.20, 2.21) used as the analogue of equation (2.41) a modified Helmholtz equation for the magnetic field strength distribution:

$$\Delta\tilde{H} + k^2\tilde{H} = -2ik \frac{\partial}{\partial x} [\phi(x)f'(x)\delta(y-f(x))]+2ik \frac{\partial}{\partial y} [\phi(x)\delta(y-f(x))] \quad . \tag{2.69}$$

Again the solution of a Fredholm equation of the first kind was necessary before the field amplitudes could be evaluated. However, in this case the use of the Fourier series method for the solution of the integral equation led to an expression involving the summation of a divergent series. This required the employment of an empirical renormalization. The results of the renormalized calculation were shown to be of comparable accuracy to those provided by the Rayleigh theory in two cases of diffraction by a triangular profile. However, the use of the renormalization could not be justified theoretically, and the limits of

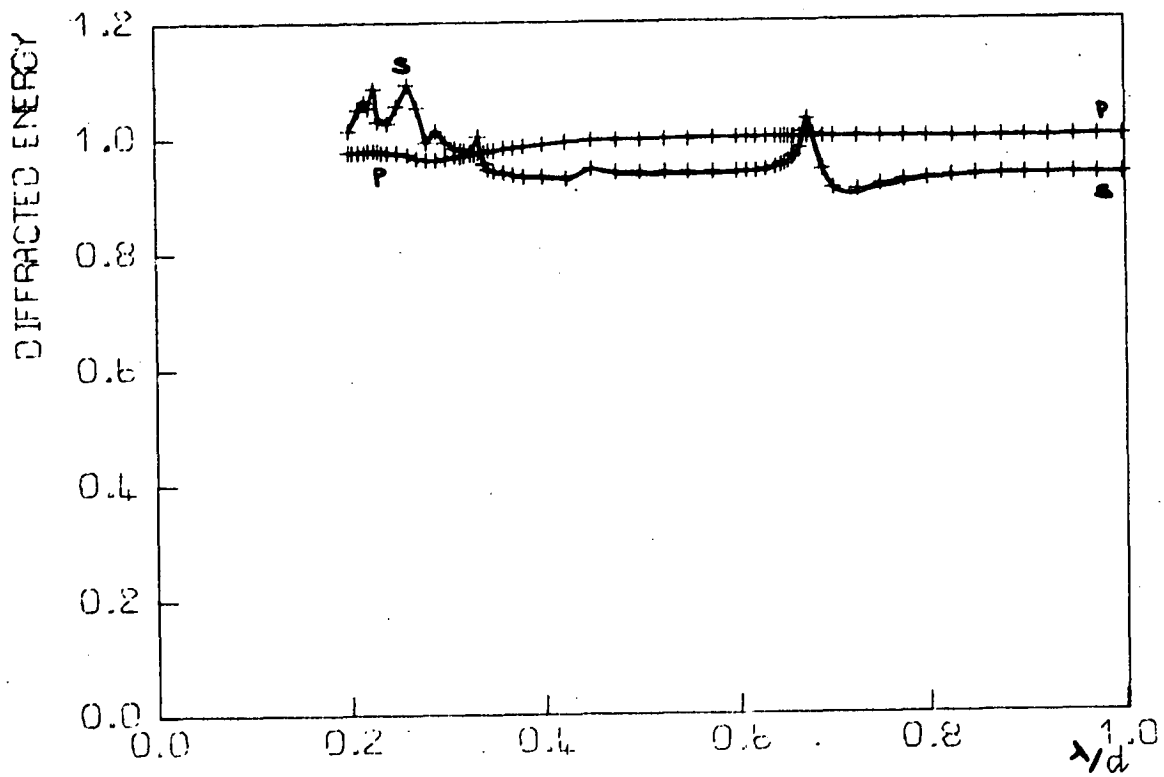


Figure (2.13). Diffracted energy curves computed using Fourier series methods, for a grating having a triangular profile with facet angles 10^{o} and 80^{o}, used at a constant angle of incidence ($\theta = 20^{o}$). Curves for both P and S polarized radiation are shown.

its region of applicability could not be established.

A formulism not requiring the use of a renormalization was devised by R. Petit and M. Cadilhac (2.22). The complex amplitude of the diffracted magnetic field strength is described by a function u(x,y) which satisfies the modified Helmholtz equation

$$\Delta u + k^2 u = 2ik\, \psi(x)\, H(f(x) - y)\ . \qquad (2.70)$$

Here $\psi(x)$ is an unknown function, while H denotes the Heaviside step function. ψ is determined from a Fredholm integral equation, which can be solved by the Fourier series technique employed in the case of P polarization. The superiority of this technique over the method of renormalization has been demonstrated (2.22).

Computer programs have been written to calculate the diffraction efficiencies provided by triangular profile gratings in P and S polarized light, based respectively on the first and third of the above formulations. These programs were used to calculate the diffracted energy curves of Figures (2.13) and (2.14). By comparing these curves with those of Figures (2.9) and (2.10), an idea of the relative powers of the Fourier series and Rayleigh methods for the case of the triangular profile can be gained.

Figure (2.13) corresponds to a grating with facet angles 10^o and 80^o, used at an angle of incidence of 20^o. For P polarization, the efficiency calculations were made with Q = 4 and M = 9. The diffracted energy varies smoothly with wavelength, just as did the corresponding Rayleigh theory curve. The two theories are of equivalent accuracy. For S polarization, the results shown are a combination of some obtained with Q = 4 and M = 9, and others obtained with Q = 6 and M = 13. The diffracted energy varies in an oscillatory and sometimes ragged fashion

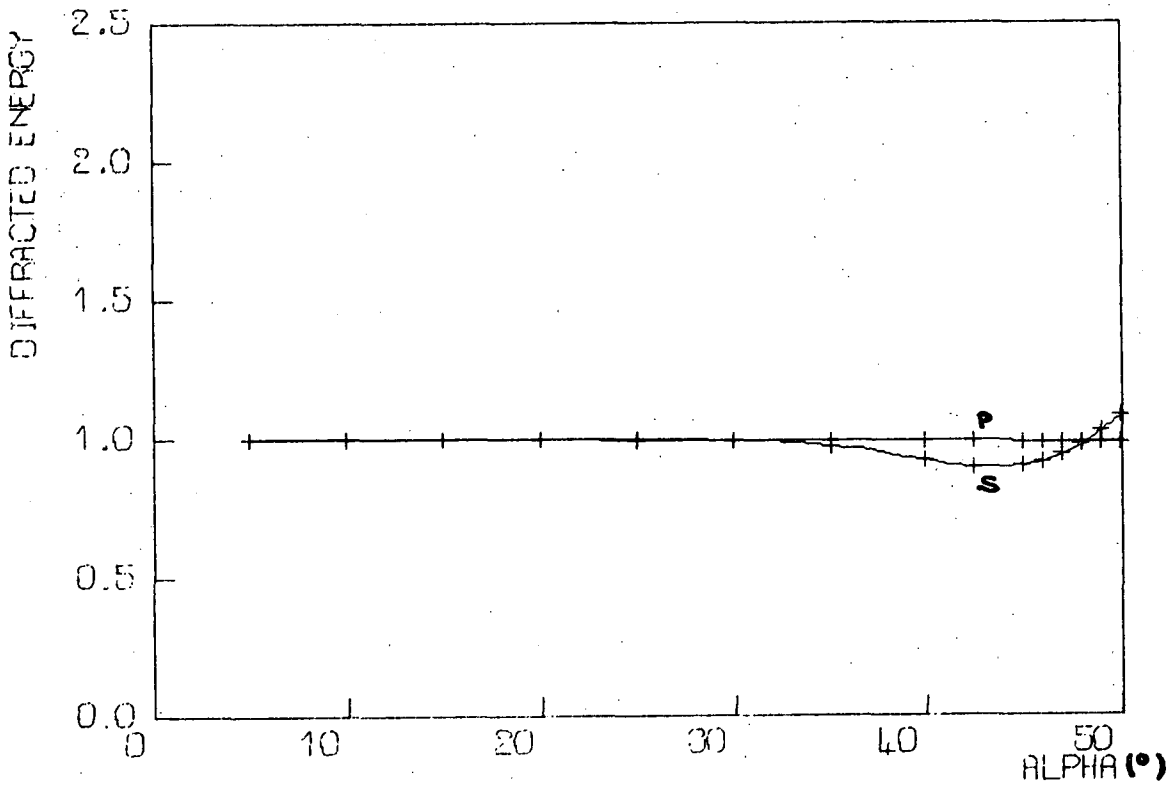


Figure (2.14). The variation of diffracted energy as a function of the blaze angle (α) in degrees is shown, for gratings having symmetrical triangular profiles, used with normally incident light of wavelength $\lambda/d = 0.4368$. (Fourier series results.)

with wavelength. Again, the accuracies of the two theories are comparable. (For the Rayleigh results, the limits of variation of the diffracted energy are 0.86 and 1.06. The corresponding limits for the Fourier series results are 0.91 and 1.09;)

Figure (2.14) shows diffracted energy curves for symmetrical ($\alpha = \beta$) triangular profiles, used with normally incident radiation such that $\lambda/d = 0.4368$. For P polarization, the diffracted energy decreases slowly from 100% as the angle α increases. For $\alpha = 50^{o}$, the energy defect (E.D. = 1.0-D.E.) is still only 1.3%. For S polarization, the accuracy is good until α passes through 30^{o}. By $\alpha = 42.5^{o}$ the energy defect has increased to 10.3%. As α increases beyond this value, the energy defect decreases, and is equal to -8.5% for $\alpha = 50^{o}$.

This test illustrates the greater accuracy of the Fourier series treatment when compared with the Rayleigh formulism, for which serious errors began to occur before $\alpha = 35^{o}$. Also, the superiority of the P polarization results over those for S polarization can be seen when the profile depth becomes comparable with the wavelength.

2.5. THE FORMULATION OF PAVAGEAU AND BOUSQUET

The description of this formulation is based on references (2.23-2.25). Diffraction by a perfectly conducting, cylindrical surface having a directrix C is considered. The incident field is specified by its magnetic field vector $\underline{H}^{i}$. This field induces a surface current on the grating surface, of density $\underline{i}$. The retarded vector potential of the diffracted wave is

$$\underline{A}(P) = -\frac{j\,\mu_0}{4} \int_C \underline{i}(M)\, H_0^{(2)}\,(k,P,M)\, dM\,, \qquad (2.71)$$

where μ_0 is the magnetic permeability of free space, and $H_0^{(2)}$ is the

Hankel function of the second kind and order zero. The diffracted magnetic field can also be obtained from $\underline{i}$:

$$\underline{H}(P) = -\frac{jk}{4} \int_C \underline{i}(M) \times \frac{\underline{MP}}{|\underline{MP}|} H_1^{(2)}(k,P,M)\, dM \quad . \qquad (2.72)$$

The fields at the surface of the grating are defined to be the average of the one-sided limits of the fields in the half-spaces above and below the surface. Thus, the boundary condition for this problem is

$$\underline{i} = 2\underline{n} \times (\underline{H}^i + \underline{H}) \quad , \qquad (2.73)$$

where $\underline{n}$ is the unit normal vector to the surface which is oriented towards the vacuum, and $\underline{H}$ is the diffracted field at a point P lying strictly on the surface. Using equation (2.72), this boundary condition gives

$$\underline{i}(P) = 2\underline{n}(P) \times [\underline{H}^i(P) - \frac{jk}{4} \oint_C \underline{i}(M) \times \frac{\underline{MP}}{|\underline{MP}|} H_1^{(2)}(k,P,M)dM \,]. \qquad (2.74)$$

The symbol $\oint_C$ denotes the principal value of the integral obtained by avoiding the singularity of the Hankel function at M = P. Equation (2.74) is one form of the basic integral equation for the problem of diffraction by a perfectly conducting grating.

Suppose now the grating has directrix z = f(x), period a and is struck by an incident wave at an angle i. Since the grating has period a, we have for any integer p,

$$\underline{i}(x+pa) = \underline{i}(x) \exp[-jk\, pa \sin i \,] \quad , \qquad (2.75)$$

so that it is sufficient to determine the values of $\underline{i}$ in one groove of the grating.

The Fourier series for $H_0^{(2)}$ takes the form

$$H_0^{(2)}(k;\, x,z\,;\, x',z') = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\exp -jk[\alpha(x-x')+\gamma|z-z'|]}{\gamma}\, d\alpha, \quad (2.76)$$

provided that $z \neq z'$. Here

$$\begin{aligned} \gamma &= \sqrt{1-\alpha^2} \quad \text{if } \alpha^2 \leq 1 \\ &= -j\sqrt{\alpha^2-1} \quad \text{if } \alpha^2 > 1\,. \end{aligned} \quad (2.77)$$

Using equations (2.75) and (2.76), the formula (2.71) becomes

$$\underline{A}(P) = -\frac{j\mu_0}{4\pi} \sum_{p=-\infty}^{\infty} \oint_0^a \underline{i}(x')\sqrt{1+f'^2(x')} \int_{-\infty}^{\infty} e^{jp\,ka(\alpha-\sin i)} \frac{e^{-jk[\alpha(x-x')+\gamma|z-f(x')|]}}{\gamma}\, d\alpha dx'. \quad (2.78)$$

The principal value symbol here excludes all points such that $f(x') = z$. But, from Poisson's identity,

$$\sum_{p=-\infty}^{\infty} e^{jp\,ka(\alpha-\sin i)} = \frac{2\pi}{ka} \sum_{p=-\infty}^{\infty} \delta\left(\alpha-\sin i - \frac{2\pi p}{ka}\right). \quad (2.79)$$

Thus, if

$$\alpha_p = \sin i + \frac{2\pi p}{ka} \quad (2.80)$$

and

$$\begin{aligned} \gamma_p &= \sqrt{1-\alpha_p^2} \quad \text{if } \alpha_p^2 \leq 1 \\ &= -j\sqrt{\alpha_p^2-1} \quad \text{if } \alpha_p^2 > 1 \end{aligned} \quad (2.81)$$

then

$$\underline{A}(P) = -\frac{j\mu_0}{2ka} \oint_0^a \underline{i}(x')\sqrt{1+f'^2(x')} \sum_{p=-\infty}^{\infty} \frac{e^{-jk[\alpha_p(x-x')+\gamma_p|z-f(x')|]}}{\gamma_p}\, dx'. \quad (2.82)$$

The corresponding formula for the diffracted magnetic field is

$$\underline{H}(P) = \frac{1}{2a} \oint_0^a \sqrt{1+f'^2(x')}\, \underline{i}(x') \times \sum_{p=-\infty}^{\infty} \left(\frac{\alpha_p}{\gamma_p} \underline{u}_x + s\, \underline{u}_z \right)$$

$$e^{-jk[\alpha_p(x-x')+\gamma_p|z-f(x')|]}dx', \qquad (2.83)$$

where $\underline{u}_x$ and $\underline{u}_z$ are the unit vectors along the Ox and Oz axes, and $s = \text{sign}(z-f(x'))$.

If $z > \max[f(x')]$, the above formula reduces to the plane-wave expansion

$$\underline{H}(P) = \sum_{p=-\infty}^{\infty} \underline{H}_p\, e^{-jk(\alpha_p x+\gamma_p z)} \quad , \qquad (2.84)$$

where

$$\underline{H}_p = -\frac{1}{2a} \left(\frac{\alpha_p}{\gamma_p} \underline{u}_x + \underline{u}_z \right) \times \int_0^a \underline{i}(x') \sqrt{1+f'^2(x')}\, e^{jk[\alpha_p x' + \gamma_p f(x')]}dx'. \qquad (2.85)$$

Equation (2.85) enables the calculation of the efficiencies of the diffracted orders. As a check on the calculation, the vanishing of the total magnetic field beneath the grating surface implies that

$$\frac{1}{2a} \left(\frac{\alpha_p}{\gamma_p} \underline{u}_x - \underline{u}_z \right) \times \int_0^a \underline{i}(x') \sqrt{1+f'^2(x')}\, e^{jk[\alpha_p x' - \gamma_p f(x')]}dx' = H^i \delta_{p,0}, \qquad (2.86)$$

where $\delta_{p,0}$ is the Kronecker delta symbol.

The use of equation (2.83) in the boundary condition (2.73) leads to the following requirement on $\underline{i}(x)$:

$$\underline{i}(x) = 2\underline{n}(x) \times [\underline{H}^i(x,f(x)) + \frac{1}{2a} \oint_0^a \sqrt{1+f'^2(x')}\, \underline{i}(x') \times \sum_{p=-\infty}^{\infty} \left(\frac{\alpha_p}{\gamma_p} \underline{u}_x + s\underline{u}_z \right).$$

$$\exp -jk\{\alpha_p(x-x')+\gamma_p|f(x)-f(x')|\}dx']. \qquad (2.87)$$

$\underline{i}(x)$ is thus determined by a non-homogeneous Fredholm integral equation. Expressions (2.85-2.87) comprise the vectorial diffraction formulation, valid for any polarization of the incident field.

If the incident field is P polarized, then the vectorial equations can be reduced to the following expressions: if

$$\phi(x) = i(x)\sqrt{1+f'^2(x)}\ e^{jkx \sin i} \qquad (2.88)$$

and

$$\phi_0(x) = 2H^i[1 + f'(x)\sin i]\, e^{jk\, f(x)\cos i}\ , \qquad (2.89)$$

then the integral equation (2.87) becomes

$$\phi(x) = \phi_0(x) + \frac{1}{a}\int_0^a \phi(x')\, N(x,x')dx'\ , \qquad (2.90)$$

where the kernel function is

$$N(x,x') = \sum_{p=-\infty}^{\infty} [s - \frac{\alpha_p}{\gamma_p} f'(x)]\ \exp j[pK(x'-x)-k\gamma_p|f(x')-f(x)|]. \qquad (2.91)$$

The diffracted field amplitudes are given by:

$$H_p = \frac{1}{2a\gamma_p}\int_0^a \phi(x) \exp j[pKx + k\,\gamma_p\, f(x)]dx\ . \qquad (2.92)$$

The integral equation (2.90) is unaltered for the case of S polarization. The other fundamental equations become

$$\phi(x) = i(x)\, e^{jkx \sin i}\ , \qquad (2.93)$$

$$\phi_0(x) = -2H^i\, e^{jk\, f(x)\cos i}\ , \qquad (2.94)$$

$$N(x,x') = \sum_{p=-\infty}^{\infty} (s - \frac{\alpha_p}{\gamma_p} f'(x')) \exp j[pK(x'-x)-k\gamma_p|f(x')-f(x)|] \qquad (2.95)$$

and

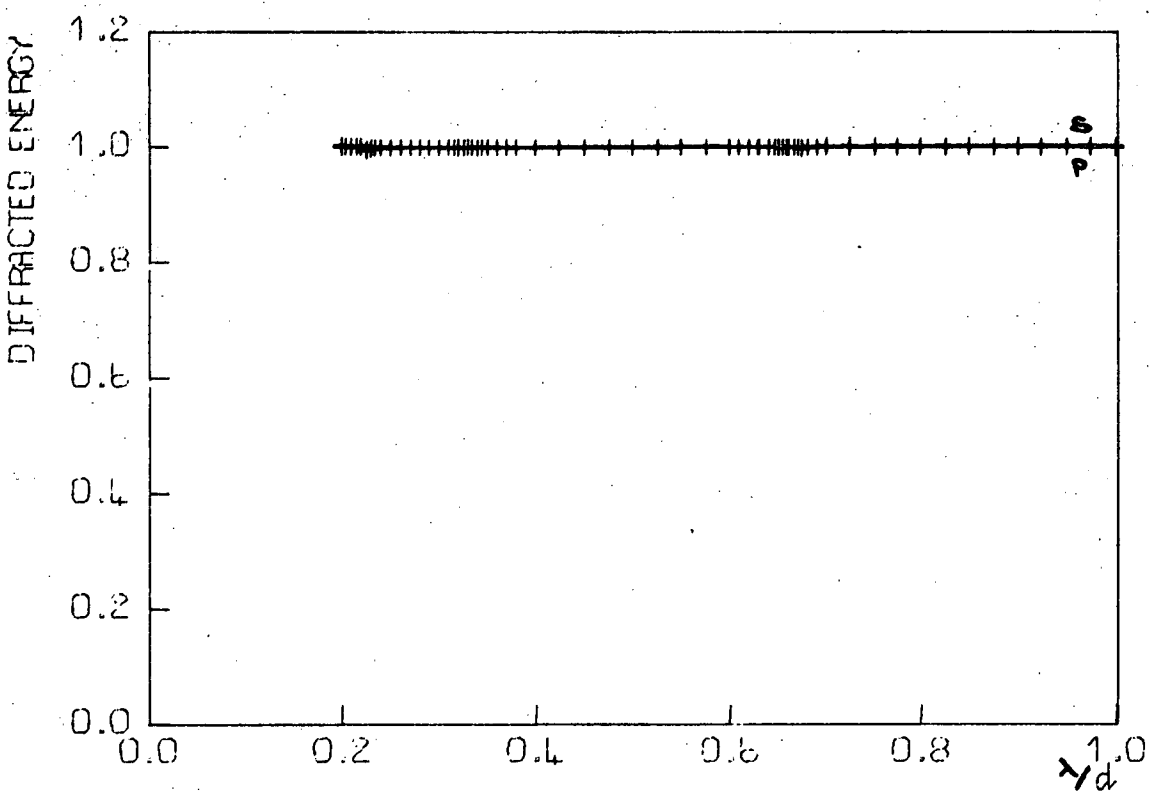


Figure (2.15). Diffracted energy curves computed using the formulism of Pavageau and Bousquet, for a grating having a triangular profile with facet angles 10^{o} and 80^{o}, used at a constant angle of incidence ($\theta = 20^{o}$). Curves for both P and S polarized radiation are shown.

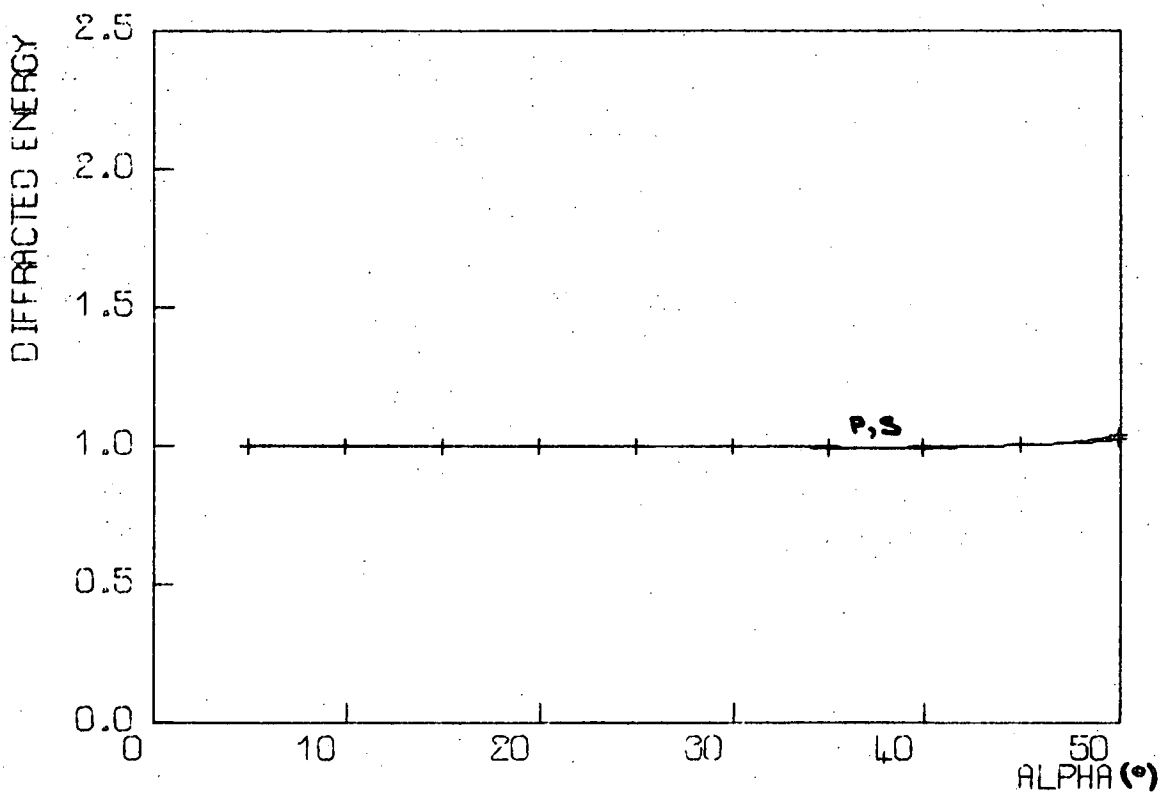


Figure (2.16). The variation of diffracted energy as a function of the blaze angle (α) in degrees is shown, for gratings having symmetrical triangular profiles, used with normally incident light of wavelength $\lambda/d = 0.4368$. (Results obtained using the formulism of Pavageau and Bousquet.)

$$H_p = \frac{1}{2a} \int_0^a \phi(x') \left(\frac{\alpha_p}{\gamma_p} f'(x') - 1 \right) \exp j[pKx' + k\gamma_p f(x')] dx' \quad . \qquad (2.96)$$

The equations (2.89-2.92) and (2.94-2.96) have served as the basis of programs to calculate diffraction grating efficiencies in P and S polarized light. Some aspects of these programs will be described in Appendix I. Theoretical efficiency values have been established for sinusoidal and triangular profiles, to enable comparisons to be made between this formulation and the three other methods described above.

The P and S polarization diffracted energy curves of Figure (2.15) correspond to a triangular profile grating with facet angles 10^0 and 80^0, used at a constant angle of incidence of 20^0. The calculations were made using a twenty-point, two-segment Gaussian quadrature rule. The two curves of Figure (2.15) vary over a smaller range than do the corresponding curves for the other formulations. For P polarization, over the whole wavelength range considered, the diffracted energy lies between the limits 1.005 and 0.997. For S polarization, the limiting values are 1.003 and 0.993.

Diffracted energy curves are given in Figure (2.16) for symmetrical triangular profiles used with normally incident light of wavelength $\lambda/d = 0.4368$. For P polarization, the diffracted energy first decreases slowly from one to reach 0.9966 at $\alpha = 40^0$, and then increases to 1.028 at 50^0. The accuracy of the calculations is very close to that of the Fourier series results. It is for S polarization that the accuracy of the formulism of Pavageau and Bousquet is clearly superior to that of the Fourier series method - the respective limits of variation are 0.995 to 1.039, and 0.897 to 1.085.

The diffracted energy curves for the sinusoidal profile shown

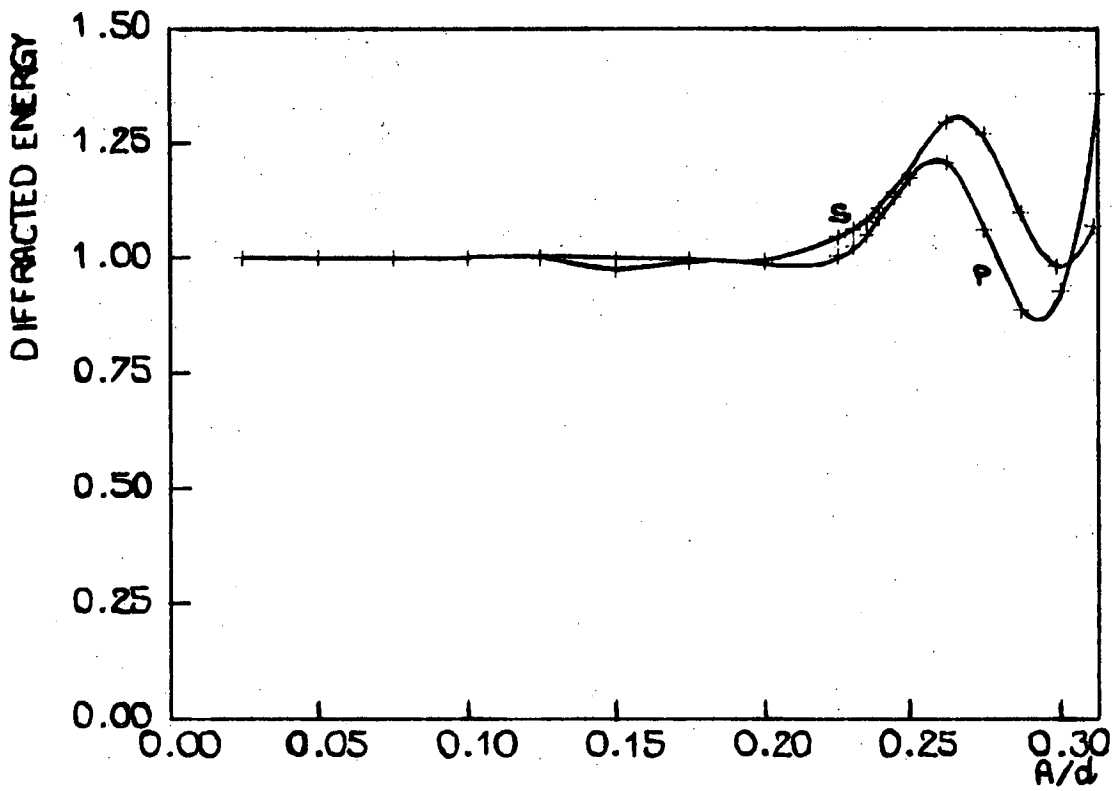


Figure (2.17). The variation of diffracted energy as a function of normalized groove amplitude is shown, for sinusoidal gratings used with normally incident light of wavelength $\lambda/d = 0.4368$. (Results obtained using the formulism of Pavageau and Bousquet.)

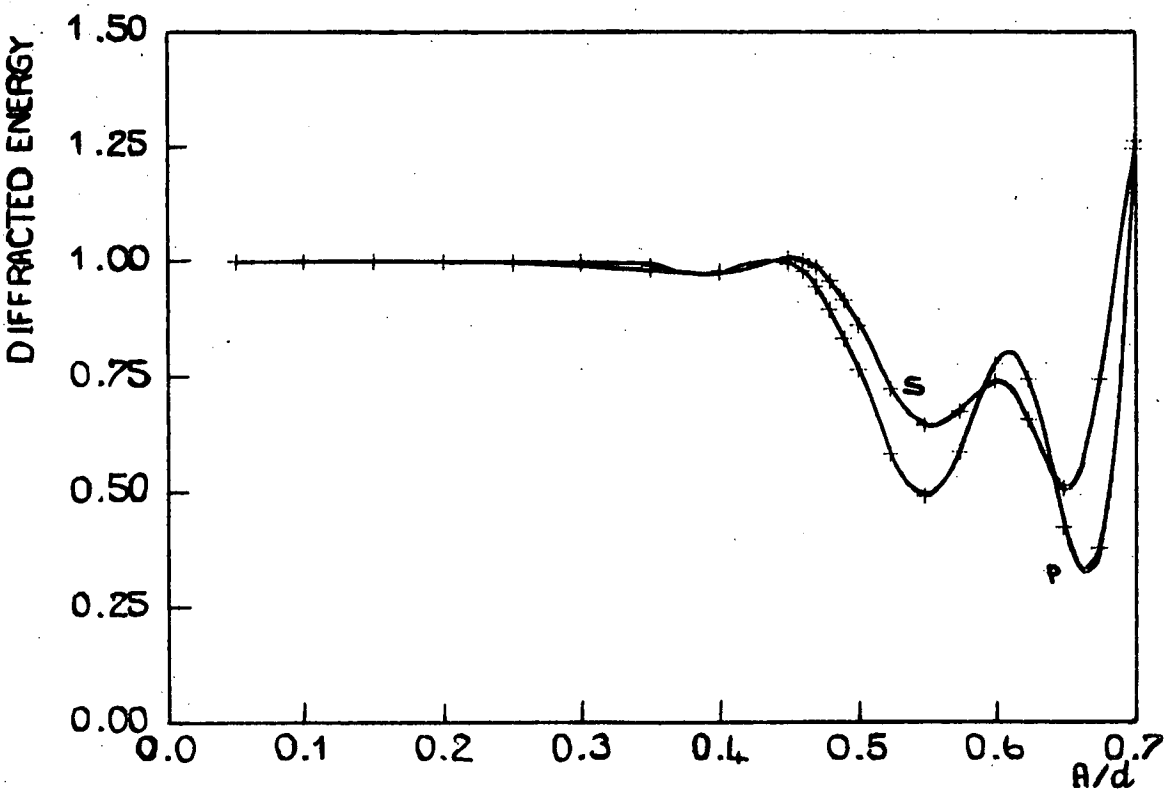


Figure (2.18). The variation of diffracted energy as a function of normalized groove amplitude is shown, for sinusoidal gratings used with normally incident light of wavelength $\lambda/d = 0.21$. (Results obtained using the formulism of Pavageau and Bousquet.)

in Figures (2.17) and (2.18) were obtained using a twenty-point Gaussian quadrature rule (as were the curves of Figure (2.16)). Comparable curves based on Rayleigh theory results are given in Figures (2.11) and (2.12).

For a normalized wavelength of 0.4368 (Figure (2.17)) the accuracy of the integral equation results is reasonable ($|E.D.| \lesssim 0.05$) for P polarization until A/d = 0.47, and for S polarization until A/d = 0.49. For the Rayleigh theory, the corresponding values are (respectively) 0.17 and 0.16.

For the second case ($\lambda/d = 0.21$) the accuracy of the integral equation results is reasonable ($|E\Phi| < 0.05$) for P polarization until A/d = 0.235, and for S polarization until A/d = 0.230. These values have been approximately halved by the decrease of the normalized wavelength. For the Rayleigh theory, the corresponding value for both P and S polarizations is 0.12. Thus, the advantage of the integral equation method is still considerable, but has declined with the decrease in λ/d.

From the above comparisons, it is evident that the formulism of Pavageau and Bousquet furnishes results of greater accuracy than those provided by the scalar, Rayleigh and Fourier methods. The first-mentioned theory can also be used for efficiency calculations with general profiles. (Only the function values, together with those of the first and second derivatives, need be supplied at integration abscissae; an analytic formula for the profile function is not necessary.) While this is also the case for the scalar and Rayleigh methods, numerical difficulties restrict the use of the Fourier technique to profiles composed of a small number of straight line segments.

TABLE (2.1)

Computation times required for one total solution of the diffraction problem on an Elliott 503 computer by various theoretical methods.

THEORY	CALCULATION PARAMETERS	COMPUTATION TIME (Minutes)
RAYLEIGH	TRIANGULAR PROFILE; L=30	0.9
	SINUSOIDAL PROFILE; L=30	4.0
FOURIER	TRIANGULAR PROFILE	
	Q=4 M=9	4.3
	Q=5 M=11	7.6
	Q=6 M=13	12.8
PAVAGEAU and BOUSQUET	GENERAL PROFILE	
	m=20	1.5
	m=30	3.5
	m=40	6.0
	TRIANGULAR PROFILE	
	m=20	1.6
	m=30	3.6

A final factor which is of interest concerning the various vectorial theories is their demand on computer time. Table (2.1) gives the calculation times (in minutes) required for the joint evaluation of P and S polarization efficiencies of all propagating orders for a single wavelength and angle of incidence on an Elliott 503 computer. While the absolute magnitudes of such times will vary from computer to computer, their relative proportions are of more general significance.

Calculations made using the formulation of Pavageau and Bousquet have generally needed only twenty points per integration, with thirty points being occasionally necessary in regions of difficult convergence. From Table (2.1), it is clear that the computation times for this formulation compare favourably with those of the other two vectorial theories.

2.6 HESSEL AND OLINER'S THEORY

We now move on to the consideration of the formulism devised by Hessel and Oliner (2.31). While their analysis can not serve as the basis of a numerical calculation of the diffraction efficiencies given by a physical grating used in a specific configuration, it is very useful in understanding some aspects of the behaviour of efficiency values calculated using other methods, as will be seen in subsequent chapters.

The method of Hessel and Oliner deals with the scattering of an incident plane wave by a plane possessing a periodically modulated surface reactance. Because of its periodicity, any grating, whether constructed of infinitely or finitely conducting metal or of dielectric, has a periodic reactance function in the plane across the top of its

grooves. However, the specific form of the reactance function for a given physical grating can only be deduced when the electric and magnetic fields are known everywhere in the plane - i.e., when the problem of the diffraction by the grating has been solved. This difficulty restricts the application of the method to the consideration of general aspects of the diffraction process, rather than specific instances.

The reactance plane is assumed to extend indefinitely in the x and y directions, and to have a period d in the x-direction. Following the treatment of Hessel and Oliner (2.31), an S polarized plane wave is taken to be incident at an angle θ on the reactance plane. The spatial term of the magnetic field of the incident wave is

$$H_i(x,z) = H\, e^{j(k_s x - K_0 z)} \quad , \tag{2.97}$$

where

$$K_0 = k \cos\theta \tag{2.98}$$

and

$$k_s = k \sin\theta \quad . \tag{2.99}$$

k is the free-space wavenumber of the incident field.

The total field satisfies at the grating plane z=0 the impedance boundary condition

$$Z^s(x) = -\frac{E_x(x,0)}{H_y(x,0)} \quad . \tag{2.100}$$

Here $Z^s(x)$ is a periodic reactance function which is assumed to be representable by the Fourier series

$$Z^s(x) = \sum_{\nu=-\infty}^{\infty} Z_\nu^s \; e^{(2\pi i \nu x/d)} \tag{2.101}$$

where

$$Z_{-\nu}^s = -Z_\nu^{s*} \quad . \tag{2.102}$$

If I_n denotes the amplitude of the nth spectral order, the far-field representation of the scattered magnetic field strength takes the usual form:

$$H_s(x,z) = \sum_{n=-\infty}^{\infty} I_n \exp i[(k_s + 2\pi n/d)x + K_n z] \qquad (2.103)$$

where

$$\begin{aligned} K_n &= [k^2-(k_s + 2\pi n/d)^2]^{\frac{1}{2}} \text{ for } k^2 \geq (k_s + 2\pi n/d)^2 \\ &= i[(k_s + 2\pi n/d)^2 - k^2]^{\frac{1}{2}} \text{ for } k^2 \ll (k_s + 2\pi n/d)^2 \end{aligned} \qquad (2.104)$$

The boundary condition (2.100) can be written as

$$\left[Z^s(x)\, H_y(x,z) + \frac{1}{i\omega\varepsilon} \frac{\partial H_y(x,z)}{\partial z} \right]_{z=0} = 0 \,. \qquad (2.105)$$

Using the series expansions (2.101) and (2.103), equation (2.105) becomes

$$\sum_{n=-\infty}^{\infty} \frac{K_n}{\omega\varepsilon} \hat{I}_n e^{2\pi i n x/d} - \frac{2H K_0}{\omega\varepsilon} + \sum_{m=-\infty}^{\infty} \sum_{\nu=-\infty}^{\infty} Z_\nu^s \hat{I}_m e^{2\pi i(\nu+m)x/d} = 0 \qquad (2.106)$$

if

$$\hat{I}_m = I_m + H \delta_{m,0} \,. \qquad (2.107)$$

Since the harmonic functions are linearly independent over the period [0,d], equation (2.106) implies that

$$\sum_{m=-\infty}^{\infty} (Z_{n-m}^s + \frac{K_n}{\omega\varepsilon} \delta_{n,m}) \hat{I}_m = \frac{2K_0}{\omega\varepsilon} H \delta_{n,0} \,, \qquad (2.108)$$

or in matrix form

$$(Z)\, \underline{\hat{I}} = \underline{V} \,. \qquad (2.109)$$

The amplitudes $\hat{I}_m$ are then given by

$$\hat{I}_m = \frac{\Delta_m}{\Delta} \,, \qquad (2.110)$$

where Δ is the determinant of (Z), and Δ_m is the determinant of the matrix obtained by replacing the mth column of Δ by the column vector $\underline{V}$.

Diffraction anomalies manifest themselves as rapid changes in the amplitudes $\hat{I}_m$ with wavelength or angle of incidence. One type of anomaly is associated with the condition

$$K_n = [k^2 - (k_s + \frac{2\pi n}{d})^2]^{\frac{1}{2}} = 0 \ . \qquad (2.111)$$

Equation (2.111) corresponds to the occurrence of a branch-point singularity in the functional dependence of $\hat{I}_m$ on k or k_s, because of the change in form of K_n at its zero expressed in equation (2.104). Also, the derivative of K_n with respect to k or k_s becomes infinite when equation (2.111) is satisfied. Both the branch-point and steep-gradient aspects of these S polarization Wood anomalies will be discussed further in the next chapter.

A second type of anomalous behaviour is associated with the vanishing of the determinant Δ, and the condition

$$D_n = Z_0^s + \frac{K_n}{\omega\epsilon} = 0 \ . \qquad (2.112)$$

Equation (2.112) implies that this new type of anomaly, called by Hessel and Oliner (2.31) the class of resonance anomalies, will occur at a wavelength or angle of incidence distinct from that at which the nth propagating order passes off. Near a resonance anomaly, k_s approaches a zero of the determinant Δ, so that the value of $\hat{I}_m$ becomes large and rapidly varying. The zero corresponds to a guided wave supportable by the grating, and is complex. As a consequence, k_s never becomes exactly equal to the zero, and hence $\hat{I}_m$ always remains finite.

For a metallic grating with shallow grooves, Hessel and

Oliner (2.31) show that the S polarization Wood and resonance anomalies occur near to each other and thereby enhance each other's effects. It will be seen in Chapter 3 that, as the groove depth of the grating increases, the resonance anomalies become wider and move further away from the Rayleigh wavelength. The existence of two separate classes of S polarization grating anomalies is thus shown clearly.

An analysis similar to the one described for S polarization can also be constructed for P polarization (2.31). The existence of P resonance anomalies is seen to require deeper groove depths than the shallow values which are associated with sharp S resonances. This will be confirmed in Chapter 4.

2.7. THE VERIFICATION OF RIGOROUS THEORETICAL EFFICIENCY CALCULATIONS

We have followed the development of rigorous methods of calculation of the diffraction efficiencies given by gratings. It is now appropriate to investigate the quality of the agreement between theoretical efficiency curves and reliable experimental measurements.

We have already seen in Figure (1.12) that vectorial calculations of Kalhor and Neureuther (2.33) are in good agreement with P polarization efficiency measurements made at millimetre wavelengths by Deleuil (2.34). In fact, the agreement between theory and the millimetre-wave measurements is much better than is the agreement between the latter and measurements made in the infrared by Madden and Strong (2.1) on a nominally equivalent grating. The same comment applies to the corresponding values for S polarization shown in Figure (2.19). This figure is based on Kalhor and Neureuther's Figure (9), but an extra theoretical curve has been added, based on the data of Petit (2.35).

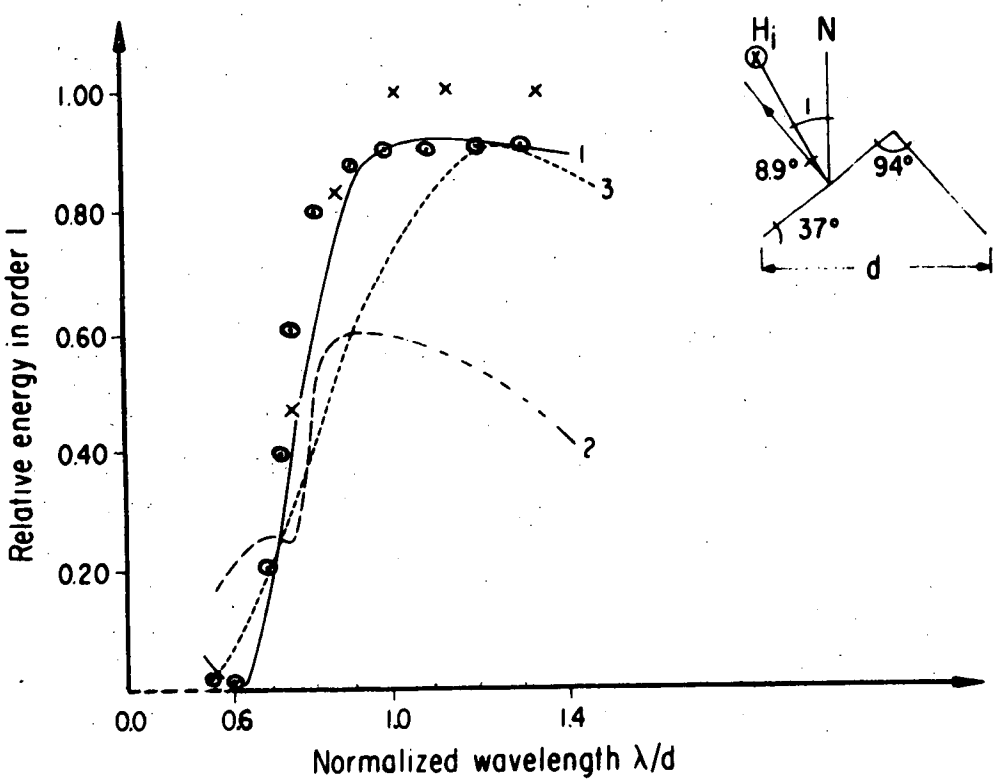


Figure (2.19). S polarization efficiency curves for an echelette grating of period d, used with a constant angular deviation of 8.9^0 in the first order. Curves 1 and 2 correspond respectively to the millimetre-wave measurements of Deleuil, and the infrared measurements of Madden and Strong. Curve 3 was obtained using the scalar theory of Madden and Strong, and the crosses and circles denote respectively the vectorial theory values of Kalhor and Neureuther and of Petit. (After Kalhor and Neureuther.)

Although Petit's values satisfy badly the criterion of conservation of energy, they are still in better agreement with Deleuil's measured values than are the calculations of Kalhor and Neureuther, particularly at longer wavelengths.

Figure (2.20) is based on Kalhor and Neureuther's Figure 7, and again corresponds to S polarization diffraction by triangular profile gratings. Both vectorial theories, and even scalar calculations, give results in better agreement with the millimetre-wave measurements than are the infrared efficiency values measured by Madden and Strong (2.1). It is of interest that Kalhor and Neureuther's technique does not yield the anomalous behaviour near $\lambda/d = 0.7$ seen in the curves of both Petit and Deleuil.

We now turn to the comparisons of theory and experiment which have been made with lamellar gratings. The profile functions associated with such gratings have infinite derivatives at the edges of each rectangular protrusion. For this reason, the various methods of solution of the diffraction problem discussed previously in this chapter break down for this type of grating.

Theoretical formulisms which deal specifically with the diffraction by lamellar gratings have been devised by A. Wirgin (2.36-2.40). The most general formulations use two Green's functions in the derivation of expansions for the field strength above the grating and within its grooves. The application of equations of continuity to the two expansions leads to an infinite set of linear equations, which can be solved by a method of truncation to yield the amplitudes of the components of the diffracted field.

Wirgin (2.37) compared theoretical efficiency curves for S polarization with the millimetre-wave measurements of C.H. Palmer,

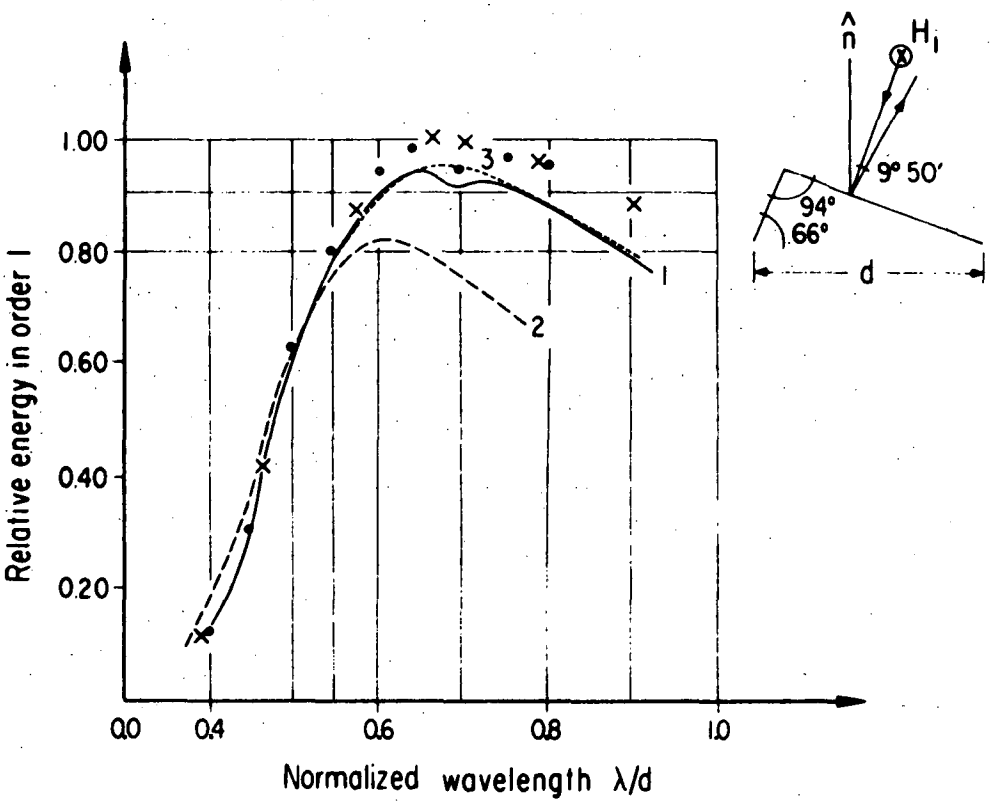


Figure (2.20). S polarization efficiency curves for an echelette grating of period d, used with a constant angular deviation of $9^{o}50'$ in the first order. Curves 1 and 2 correspond respectively to the millimetre-wave measurements of Deleuil, and the infrared measurements of Madden and Strong. Curve 3 was obtained using the scalar theory of Madden and Strong, and the crosses and dots denote respectively the vectorial theory values of Kalhor and Neureuther, and of Petit. (After Kalhor and Neureuther.)

F.C. Evering and F.M. Nelson (2.41). The discrepancies between theory and experiment are in some cases quite marked.

R. Deleuil (2.34) has also compared millimetre-wave measurements with theoretical curves of Wirgin. The differences between the two are sometimes marked. This is due to the great sensitivity of the diffraction efficiency to small changes of profile. A variation of groove depth of smaller than 3% was seen to cause a relative variation of efficiency of the order of 30%. Thus, great care must be exercised in the formation of the profiles of gratings, if measured efficiency values are to provide a meaningful test of a theory of diffraction.

Good agreement between theory and experiment has been obtained by R. Deleuil and F. Varnier (2.42), and by A. Wirgin and R. Deleuil (2.40). Figure (2.21) is taken from the latter paper. The theoretical and experimental efficiency curves shown for P polarized radiation agree everywhere to within the measurement accuracy. For S polarization the theory predicts accurately the detailed form of the zeroth order efficiency curve. The only significant departure between the two curves occurs near the peak of the bright Wood anomaly, where the corresponding theoretical and measured values are respectively 100% and 82.5%. The critical dependence of the strength of Wood anomalies on the parameters of the grating profile is clearly shown by the two S polarization curves of Figure (2.21), as well as by other curves given by Wirgin and Deleuil (2.40). Their measurements indicate that the grooves of a lamellar grating should be formed accurately to within a tolerance of $\lambda/200$, if good agreement is to be achieved between its diffraction efficiencies and theoretical calculations.

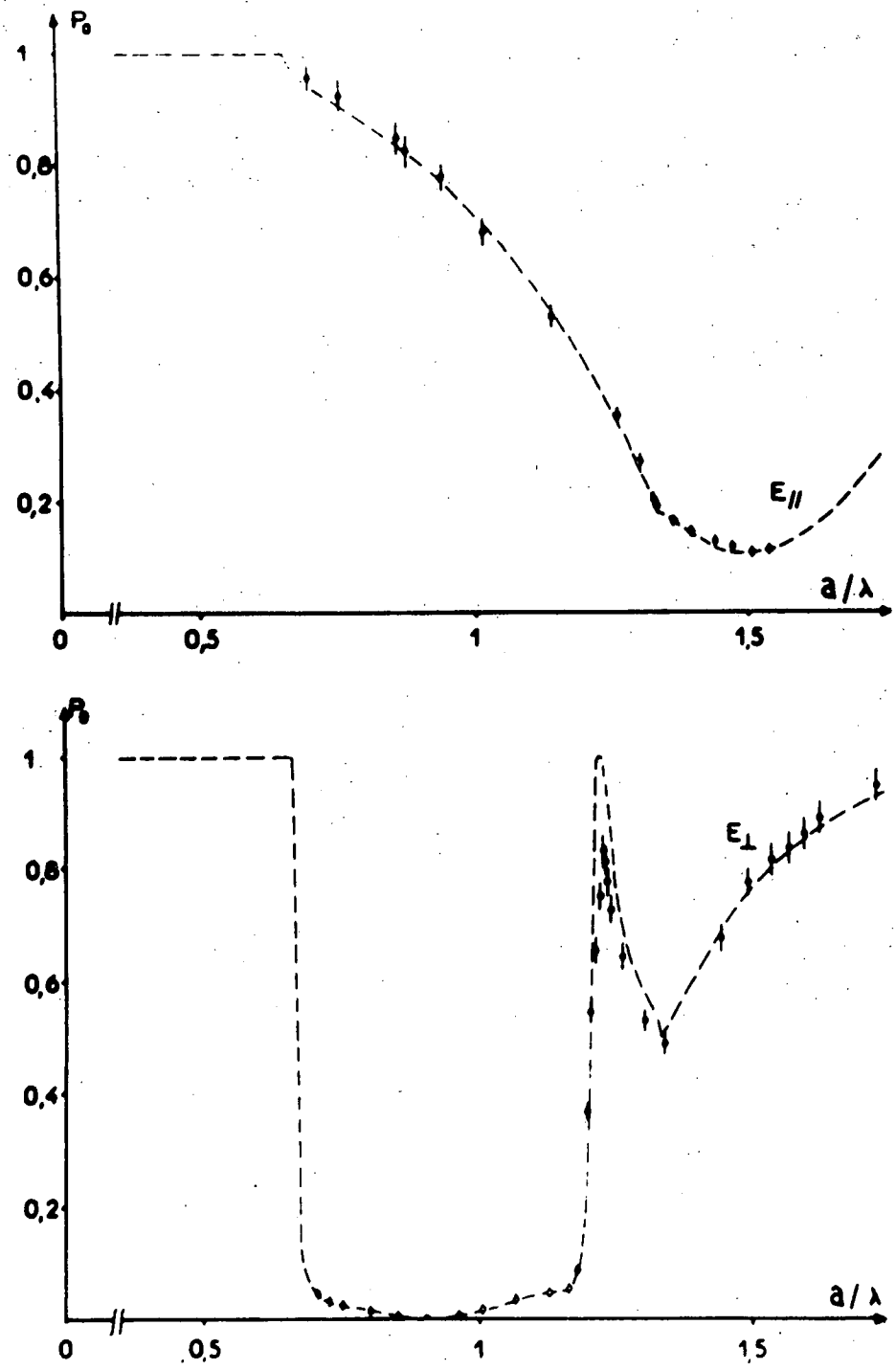


Figure (2.21). Efficiency curves for a lamellar grating. The dashed lines indicate the theoretical curves, while the points correspond to values measured at millimetre wavelengths. The top and bottom graphs refer respectively to P and S polarized radiation. (After Wirgin and Deleuil.)

It is difficult to achieve profiles of gratings to be used in the visible region which are shaped accurately to within a tolerance of the order of $\lambda/200$. Hence, comparisons of efficiencies measured at visible wavelengths and rigorous theoretical calculations have only rarely been made in the past. One such comparison was carried out by J. Cordelle, J.C. Laude, R. Petit and G. Pieuchard (2.43), with the results shown in Figure (2.22). The forms of the theoretical and experimental efficiency curves of the triangular profile grating are in good agreement, for both P and S polarizations. However, both experimental curves peak at lower efficiency values than do the corresponding theoretical curves. As well, the theoretical S polarization Wood anomaly at 0.546 μm is stronger than the observed anomaly.

Efficiency curves for four reflection gratings having triangular grooves have been determined in the visible region by G.W. Stroke (2.44). However, these curves were obtained only for unpolarized light, and in a spectrograph of unspecified configuration. Hence, it would be difficult to establish theoretical counterparts of Stroke's measured curves.

E.A. Yakovlev and F.M. Gerasimov (2.11) have made efficiency measurements on forty gratings in the visible and near infrared regions. However, precise data on the groove form is not available for these gratings. Nevertheless, R. Petit (2.35) has obtained agreement between theory and experimental measurements of the wavelength shift due to polarization made by Yakovlev and Gerasimov on a grating with a blaze angle of about 24^{o}. Yakovlev and Gerasimov's measurements will be referred to again in Chapter 5.

It is clear from the figures referred to in this section,

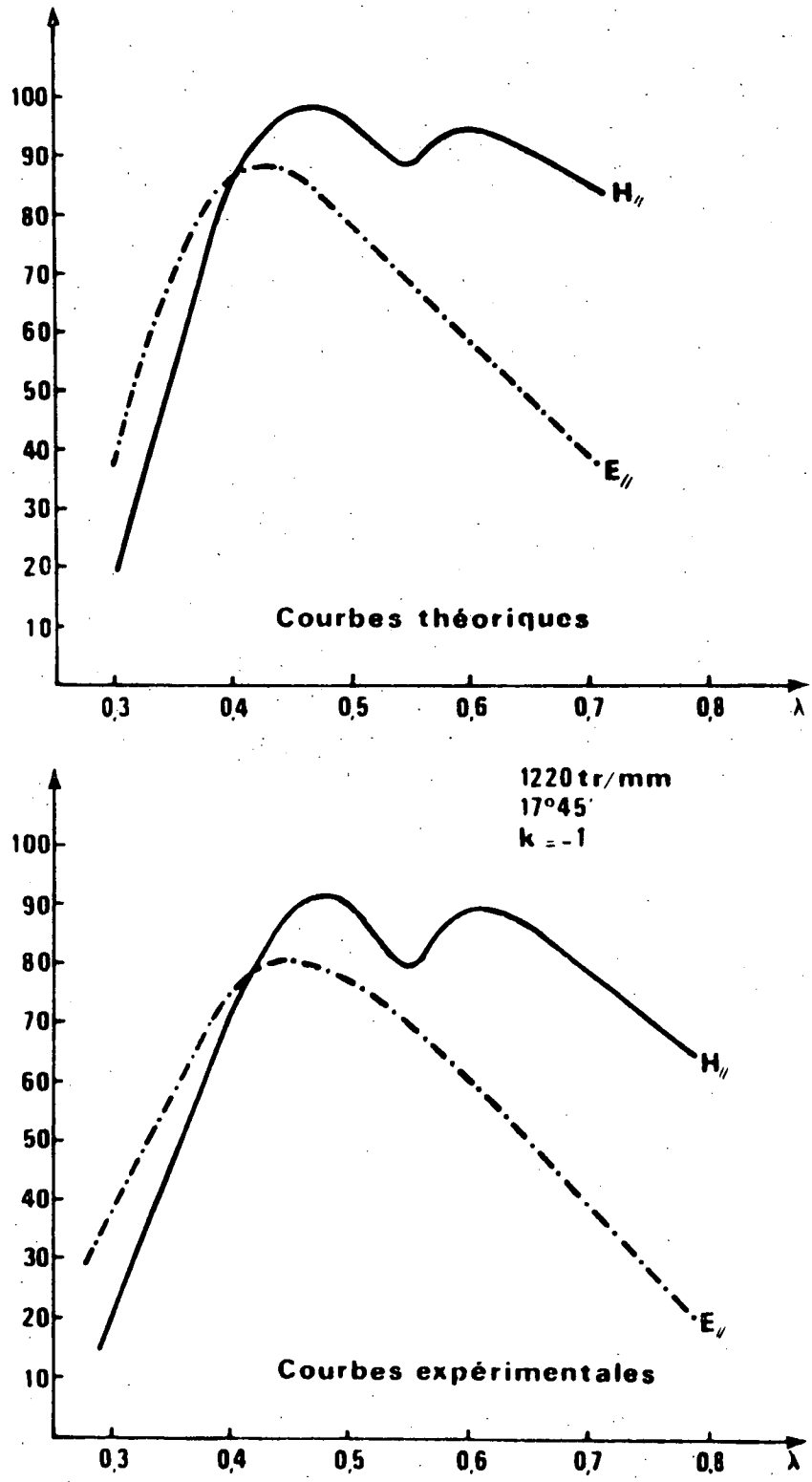


Figure (2.22). Efficiency curves for a triangular profile grating having a blaze angle of $17^{o}45'$ and a line density of 1220/mm, used in a Littrow mounting in the order -1. The top and bottom graphs have been obtained respectively by computation and by experimental efficiency measurements. (After Cordelle, Laude, Petit and Pieuchard.)

and from the considerations outlined in Section (1.5), that the most appropriate spectral region in which to test predictions based on infinite-conductivity diffraction theories is the millimetre region. The comparisons of theory and experiment made in this frequency range and displayed in Figures (2.19) to (2.21) indicate that rigorous diffraction formulisms can yield efficiency values of considerable practical importance. The two theories used in these figures are evidently capable of describing accurately diffraction phenomena given by gratings in the millimetre-wave region.

However, we cannot afford to neglect the properties of gratings which are to be used in the visible region. A problem which remains to be solved is that of determining the quality of the agreement which can be expected between the predictions of a rigorous, infinite-conductivity theory and the behaviour of optical gratings. This problem will be investigated later in the thesis.

REFERENCES.

(2.1) Madden, R.P., and Strong, J., 1958, in "Concepts of Classical Optics" (San Francisco: W.H. Freeman and Company).

(2.2) Silver, S., 1962, J.opt.Soc.Amer., 52, 131.

(2.3) Jones, D.S., 1964, "The Theory of Electromagnetism", 640 (Oxford: Pergamon Press).

(2.4) Cerutti-Maori, G., and Petit, R., 1970, Nouv.Rev.d'Opt.appl., 1, 321.

(2.5) Janot, C., and Hadni, A., 1962, J.Physique Rad., 23, 152.

(2.6) Palmer, C.H., and Le Brun, H., 1972, Appl.Optics, 11, 907.

(2.7) Uretsky, J.L., 1965, Ann.Phys., 33, 400.

(2.8) Aas, J.A., 1972, Appl.Optics, 11, 1579.

(2.9) Meecham, W.C., 1956, J.Appl.Phys., 27, 361.

(2.10) Maréchal, A., and Stroke, G.W., 1959, C.R.Acad.Sci.Paris, 249, 2042.

(2.11) Yakovlev, E.A., and Gerasimov, F.M., 1961, Optics and Spectroscopy, 10, 50.

(2.12) Petit, R., 1963, C.R. Acad. Sci. Paris, 257, 2018.

(2.13) Petit, R., 1963, Rev.d'Opt., 42, 263.

(2.14) Petit, R., 1965, Appl.Optics, 4, 1551.

(2.15) Courant, R., and Hilbert, D., 1966, "Methods of Mathematical Physics", I, 50 (New York: Interscience).

(2.16) Abramowitz, M., and Stegun, I.A., 1965, "Handbook of Mathematical Functions", 890 (New York: Dover).

(2.17) Petit, R., and Cadilhac, M., 1966, C.R.Acad.Sci.Paris, 262, 468.

(2.18) Petit, R., and Cadilhac, M., 1964, C.R.Acad.Sci.Paris, 259, 2077.

(2.19) Petit, R., 1965, C.R.Acad.Sci.Paris, 260, 4454.

(2.20) Petit, R., 1966, Rev.d'Opt., 45, 249.

(2.21) Petit, R., 1965, C.R.Acad.Sci.Paris, 261, 4677.

(2.22) Petit, R., and Cadilhac, M., 1967, C.R.Acad.Sci.Paris, 264, 1441.

(2.23) Pavageau, J., Eido, R., and Kobeissé, H., 1967, C.R.Acad.Sci. Paris, 264, 424.

(2.24) Pavageau, J., and Bousquet, J., 1969, C.R.Acad.Sci.Paris, 268, 776.

(2.25) Pavageau, J., and Bousquet, J., 1970, Optica Acta, 17, 469.

(2.26) Edwards, J., 1922, "A Treatise on the Integral Calculus", II, 194 (London: Macmillan).

(2.27) Abramowitz, M., and Stegun, I.A., 1965, "Handbook of Mathematical Functions", 16 (New York: Dover).

(2.28) Beam, A., 1960, C.A.C.M. Algorithm 14.

(2.29) Maystre, D., and Petit, R., 1970, Opt.Commun., 2, 309.

(2.30) Ralston, A., 1965, "A First Course in Numerical Analysis", 87 (New York: McGraw-Hill).

(2.31) Hessel, A., and Oliner, A.A., 1965, Appl.Optics, 4, 1275.

(2.32) Jones, D.S., 1964, "The Theory of Electromagnetism", 569 (Oxford: Pergamon Press).

(2.33) Kalhor, H., and Neureuther, A., 1971, J.opt.Soc.Amer., 61, 43.

(2.34) Deleuil, R., 1969, Optica Acta, 16, 23.

(2.35) Petit, R., 1966, Rev.d'Opt., 45, 353.

(2.36) Wirgin, A., 1966, C.R.Acad.Sci.Paris, 262, 385.

(2.37) Wirgin, A., 1966, C.R.Acad.Sci.Paris, 262, 579.

(2.38) Wirgin, A., 1966, C.R.Acad.Sci.Paris, 262, 870.

(2.39) Wirgin, A., 1966, C.R.Acad.Sci.Paris, 262, 1032.

(2.40) Wirgin, A., and Deleuil, R., 1969, J.opt.Soc.Amer., 59, 1348.

(2.41) Palmer, C.H., Evering, F.C., and Nelson, F.M., 1965, Appl.Optics, 4, 1271.

(2.42) Deleuil, R., and Varnier, F., 1968, C.R.Acad.Sci.Paris, 267, 1074.

(2.43) Cordelle, J., Laude, J.P., Petit, R., and Pieuchard, G., 1970, Nouv.Rev.d'Opt.appl., 1, 149.

(2.44) Stroke, G.W., 1964, in "Quantum Electronics III", 1221 (New York: Columbia Univ. Press).

CHAPTER 3

SOME THEORETICAL PROPERTIES OF S-ANOMALIES IN DIFFRACTION GRATINGS

This chapter is based on a paper published in Optica Acta by the author and Dr. M.D. Waterworth (3.1).

3.1. INTRODUCTION

Diffraction anomalies have been the subject of many theoretical and experimental investigations since Wood (3.2) first reported the existence of abnormal dark and light bands in the spectra of optical reflection gratings. A substantial amount of experimental information about these anomalies has been obtained, but theoretical explanations of many of these effects are lacking. (Historical surveys of theoretical and experimental research into grating anomalies may be found in papers by Twersky (3.3), Hessel and Oliner (3.4) and Stewart and Gallaway (3.5).)

It is our purpose here to demonstrate the application of a general theory of diffraction by gratings to produce results which exhibit the properties of anomalies. These results will be shown, in some cases, to agree extremely well with experimental observations. This agreement will demonstrate that the 'anomalous' behaviour of gratings is in fact as amenable to theoretical explanation as is their behaviour over wide wavelength regions. Hence no special theoretical treatments of a type common in the past (3.6-3.10) for anomalous wavelength regions are necessary.

After a brief description of the theoretical method used for the calculation of grating efficiency, and also a description of a

general result useful in understanding some properties of grating anomalies, the anomalies associated with two profile forms of practical interest are considered. (In this paper we will restrict ourselves to the subject of S-polarization grating anomalies - formed when the electric vector of the incident wave is perpendicular to the rulings of the grating. The properties of anomalies formed in light of the orthogonal (P) polarization will be considered in a later paper.) Particular emphasis is placed on the extent to which the theoretical results mirror experimental observations. Suggestions are made concerning the need for future practical observations of grating anomalies.

3.2. THEORETICAL CONSIDERATIONS

The theoretical method used for the calculation of grating efficiencies has already been described in Chapter 2 and Appendix I, and will be described only briefly here.

The integral equation formulation of Pavageau and Bousquet (3.11) is employed. This assumes the surface of the grating to be infinitely conducting. The Fredholm integral equation is solved by the linear-equation technique recommended by Maystre and Petit (3.12), using a Gaussian quadrature formula which is found to be more accurate than the trapezoidal or parabolic approximations hitherto used (see Appendix I).

Using the criterion of conservation of energy, the accuracy of the theoretical calculations carried out by computer can be estimated. All the results quoted below are accurate to better than ±1 per cent.

In what follows, the important quantity is not the wavelength of the radiation, but rather the ratio of the wavelength to the grating period. Thus theoretical predictions made in this paper for

the visible and near infra-red regions may be tested experimentally in any convenient wavelength region. This freedom of choice of the wavelength region automatically implies a corresponding freedom of choice of the grating period.

In considering some of the properties of Wood anomalies, it is of assistance to use the Reciprocity Theorem (3.13, 3.14). This implies that for gratings of arbitrary profile, the diffraction efficiency in a specific order is independent of the interchange of the directions of incidence and diffraction. (This result is valid only for infinitely conducting gratings.)

Three consequences of the Reciprocity Theorem are:

1. If θ denotes the angle of incidence of light on a grating, and $Z(n)$ the diffraction efficiency in order n, then $Z(0)$ is a symmetrical function of θ about $\theta = 0$.
2. If the order n is being used in a constant deviation mounting, then reversal of the sign of the angular deviation does not alter the efficiency $Z(n)$.
3. As a corollary of property (1), rotation of the grating profile through 180° while keeping the angle of incidence constant does not affect the efficiency $Z(0)$. If only orders -1 and 0 are propagating, then, by the conservation of energy, $Z(-1)$ will be unaffected by the rotation also.

3.3. SINUSOIDAL PROFILE ANOMALIES

The behaviour of gratings having a sinusoidal profile has become of practical interest since the possibility of holographic

generation of such profiles has been demonstrated (3.15). Such gratings are also convenient to consider theoretically since their profile requires only two parameters for complete specification, namely the period and the amplitude of the sine wave. The triangular grating, however, requires three parameters to fix its profile, namely the period and the two angles between the groove facets and the plane of the grating. This profile will be considered in the next section.

3.3.1. Amplitude dependence of an anomaly

We will illustrate the effect of increasing the amplitude (A) of the sinusoidal profile on the form of a specific Wood anomaly. The grating here is described by a functional relationship of the form $f(x) = A \sin(2\pi x/d)$ where d is the profile period. The grating considered will have a period of 1 μm, and will be used in a Littrow mounting. The anomaly will be that due to the simultaneous passing-off of orders -3 and +2, leaving four propagating spectral orders in which the anomalous effects can be studied.

Figure (3.1) shows the variation of efficiency in the order +1, Z[+1], as a function of wavelength (in microns) for six values of the profile amplitude. Two curves showing the variation of intensity in the order +2, I[+2], are also given. Note the extreme sharpness of the anomaly in figure (3.1). Figure (3.1)(b) shows that a rapid increase in the intensity in the order +2 occurs just after its passing-off. (Note that the intensity associated with a given order is the square of the modulus of its wave function, and so can be non-zero when the order has passed off, whereas the efficiency, which is the ratio of the nett energy diffracted into the order to the incident wave energy, is zero after passing off.) This is due to the resonance

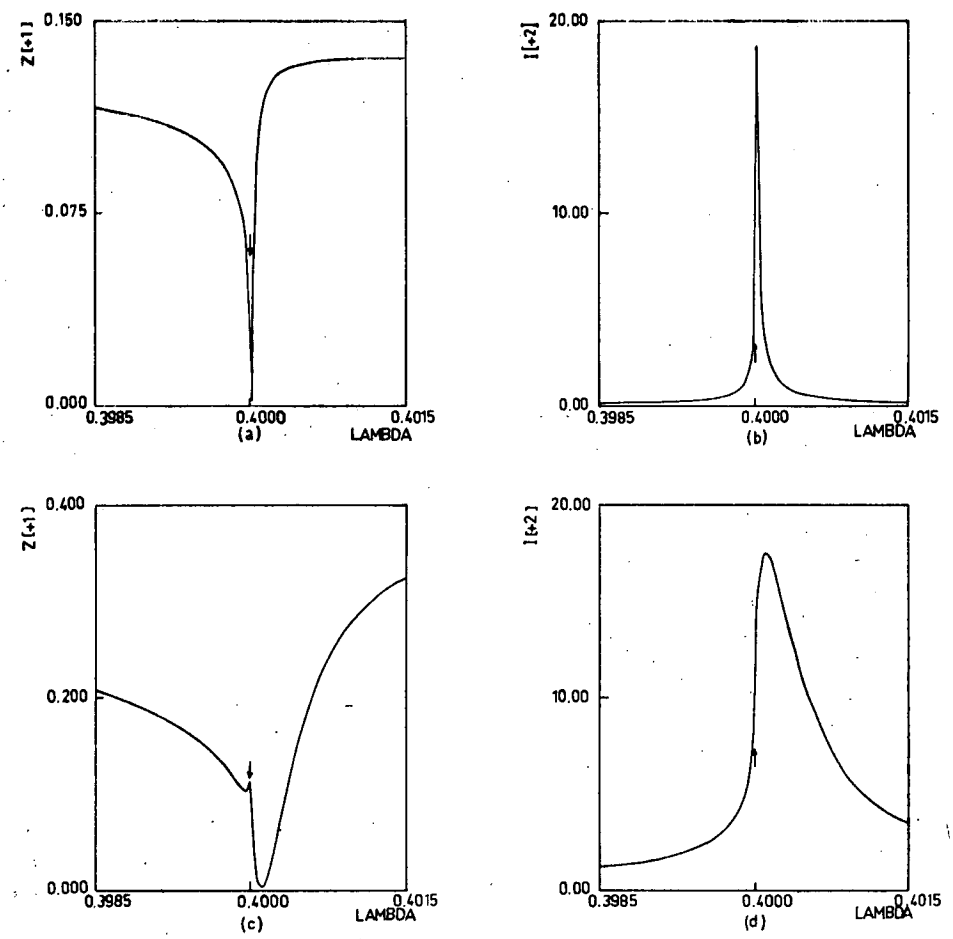


Figure 3.1 The effect of increasing amplitude on a double anomaly.

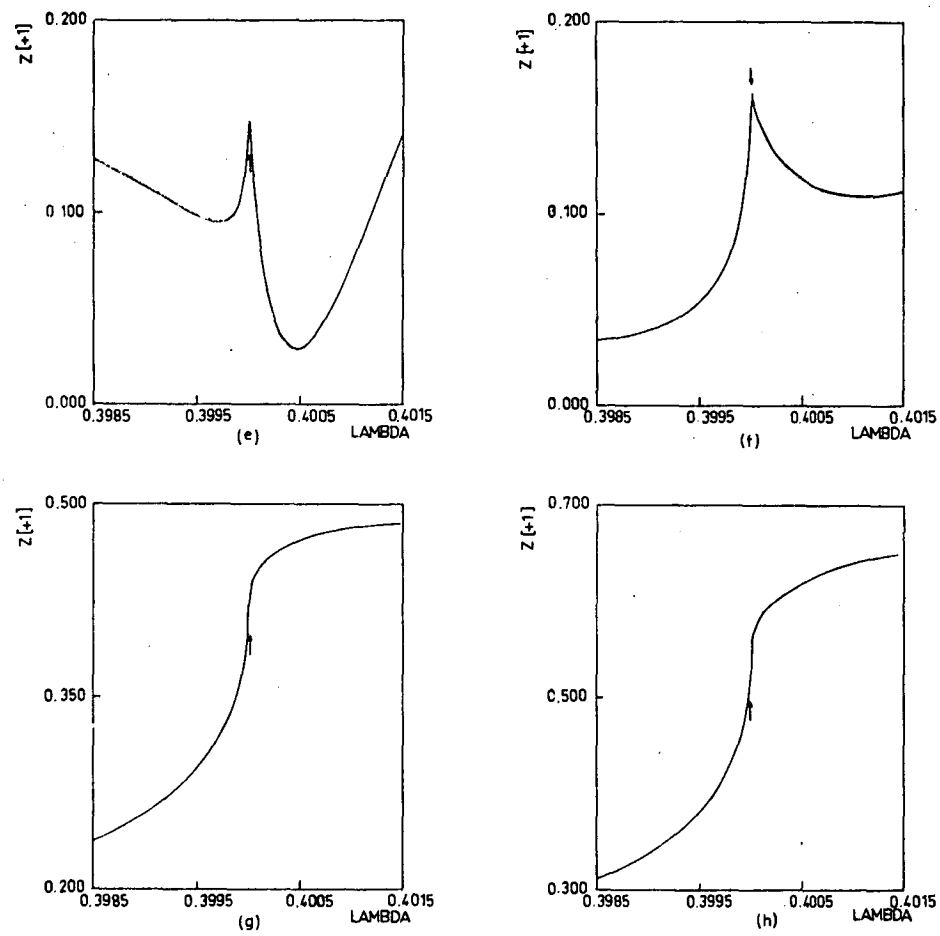
(a) A = 0.025 μm, efficiency in order +1;

(b) A = 0.025 μm, intensity in order +2;

(c) A = 0.05 μm, efficiency in order +1;

(d) A = 0.05 μm, intensity in order +2;

Figure 3.1 (continued)



(e) A = 0.075 μm,
(f) A = 0.10 μm,
(g) A = 0.15 μm,
(h) A = 0.20 μm,

efficiency in order +1

The Rayleigh wavelength is indicated by an arrow.

of this order with a complex guided wave which can be supported by the grating (3.4). The minimum in the efficiency curve occurs at the same wavelength as the maximum in the intensity curve.

When the amplitude of the profile is doubled, the bandwidth of the resonance curve increases, as does the separation between the peak of the resonance and the Rayleigh wavelength (the wavelength at which the orders -3 and +2 just graze the grating surface). The Rayleigh wavelength is now associated with a small maximum on the edge of the wide dip in the efficiency curve. Figure (3.1)(c) thus demonstrates the existence of both resonance and Wood anomalies.

The trend of increasing bandwidth and increasing separation of the resonance peak and the Rayleigh wavelength is continued as the profile depth increases up to one wavelength (figure (3.1)(h)). As a consequence of the increasing separation of the two anomalies, the influence of the resonance anomaly on the sectionsof the efficiency curve shown becomes increasingly small.

Figures (3.1)(e) and (f) show cusp-like Wood anomalies. The Rayleigh wavelength is associated with a discontinuity in the gradient of the efficiency curve. The anomaly can be regarded as being due to a branch-point in the function determining diffraction efficiency from the physical parameters of the situation (3.4). (The branch-point occurs because the cosines of the angle of diffraction of the grazing orders become imaginary as the orders pass off.)

Figures (3.1)(g) and (h) show Wood anomalies in the form of edges, with the Rayleigh wavelength corresponding to the almost vertical sections of the curves just on the short-wavelength sides of the edges.

The phenomena indicated in figures(3.1)(a) and (c) are in

keeping with the experimental observations of Ingersoll (3.16). He observed S anomalies in the form of minima, with the wavelength at the minimum always exceeding that associated with the passing-off of another order. Furthermore, Ingersoll noted that the interval between the two was largest for deeply ruled gratings. The interpretation of this behaviour in terms of resonance anomalies of the type shown in figure (3.1) is clear.

Figure (3.2) shows that a linear relationship exists between the logarithms of the half-width of the resonance anomaly and the profile amplitude, at least for profile depths smaller than one wavelength. For deeper profiles, the half-width varies less rapidly with amplitude than in the power-law region. The half-width varies approximately as the fourth power of the amplitude for groove depths smaller than a wavelength. The mean relationship for $A < 0.20$ is:

$$HW \propto A^{3.93} .$$

In table (3.1), we give the displacement $(\lambda_M - \lambda_R)$ of the peak of the resonance curve in order +2 from the Rayleigh wavelength, the half-width of the anomaly (HW) and the intrinsic resolving-power of the anomaly $(R = \lambda_R/HW)$ for nine values of the amplitude A.

3.3.2 Dependence of the form of the anomaly on the order of diffraction

Figure (3.3) gives theoretical efficiency curves illustrating the experimental observation that the forms and strengths of Wood anomalies may be entirely different in the various orders formed by a grating (3.10, 3.17, 3.18). Figure (3.3)(a) shows an anomaly which can be described as a bright line in the centre of a wide dark band, while the anomalies in figures(3.3)(b) and (c) are, respectively, a strong bright band and a dark band.

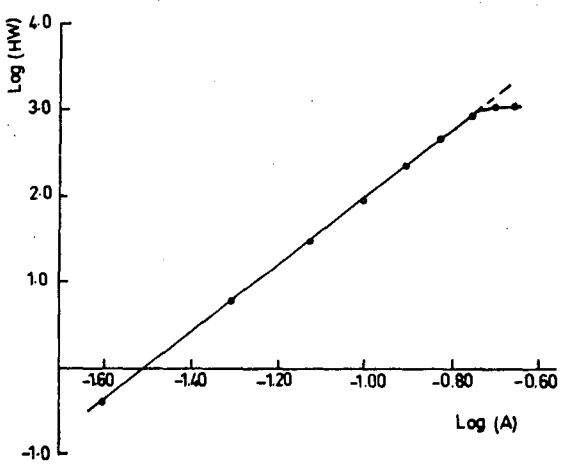


Figure 3.2. The relationship between half-width of the double anomaly and amplitude of the grating.

Table (3.1)

A(μ)	$\lambda_M - \lambda_R$ (Å)	HW(Å)	R
0.025	0.1	0.42	9520
0.050	1.1	6.10	656
0.075	4.4	29.5	136
0.100	12.5	90.0	44.4
0.125	27.0	220	18.2
0.150	78.0	484	8.26
0.175	311	859	4.66
0.200	606	1025	3.90
0.225	879	1054	3.80

Figures (3.3)(d) and (e) correspond to the double anomaly at 0.4μ due to the passing-off of orders -3 and +2. Notice the weakness of the anomaly in order -2 when compared with that in order -1. Also, the anomaly in order 0 is weak when compared with that in order +1. These facts demonstrate the complicated nature of the coupling between the propagating orders, the grazing orders and the evanescent orders formed by the grating. It is this coupling through Maxwell's equations and the boundary conditions at the grating surface that is responsible for Wood anomalies.

3.3.3. Merging of anomalies

Figure (3.4) illustrates the merging properties of the Wood anomalies associated with the grazing orders -3 and +2. For an angular deviation of zero, the theoretical positions of the two Wood anomalies coincide. As can be seen, the two Wood anomalies in all the four orders shown do combine to give single anomalies. This combination of anomalies is in agreement with the observational data of Wood (3.19) and Ingersoll (3.16), but differs from that of Stewart and Gallaway (3.5), who reported a reluctance of anomalies to merge.

Consider first the behaviour of the anomalies in the order -1 (figure (3.4)(b)). For an angular deviation of 1.5^{o}, both anomalies are edges, with the shorter wavelength one being steeper and stronger. This is also true for an angular deviation of -1.5^{o}. Thus the anomalies are not affected by the change of sign of the angular deviation, even though this change reverses the order in which the grazing orders pass off. This means that the anomaly due to the passing-off of the order -3 for an angular deviation of -1.5^{o} corresponds to that due to the passing-off of the order +2 for an angular

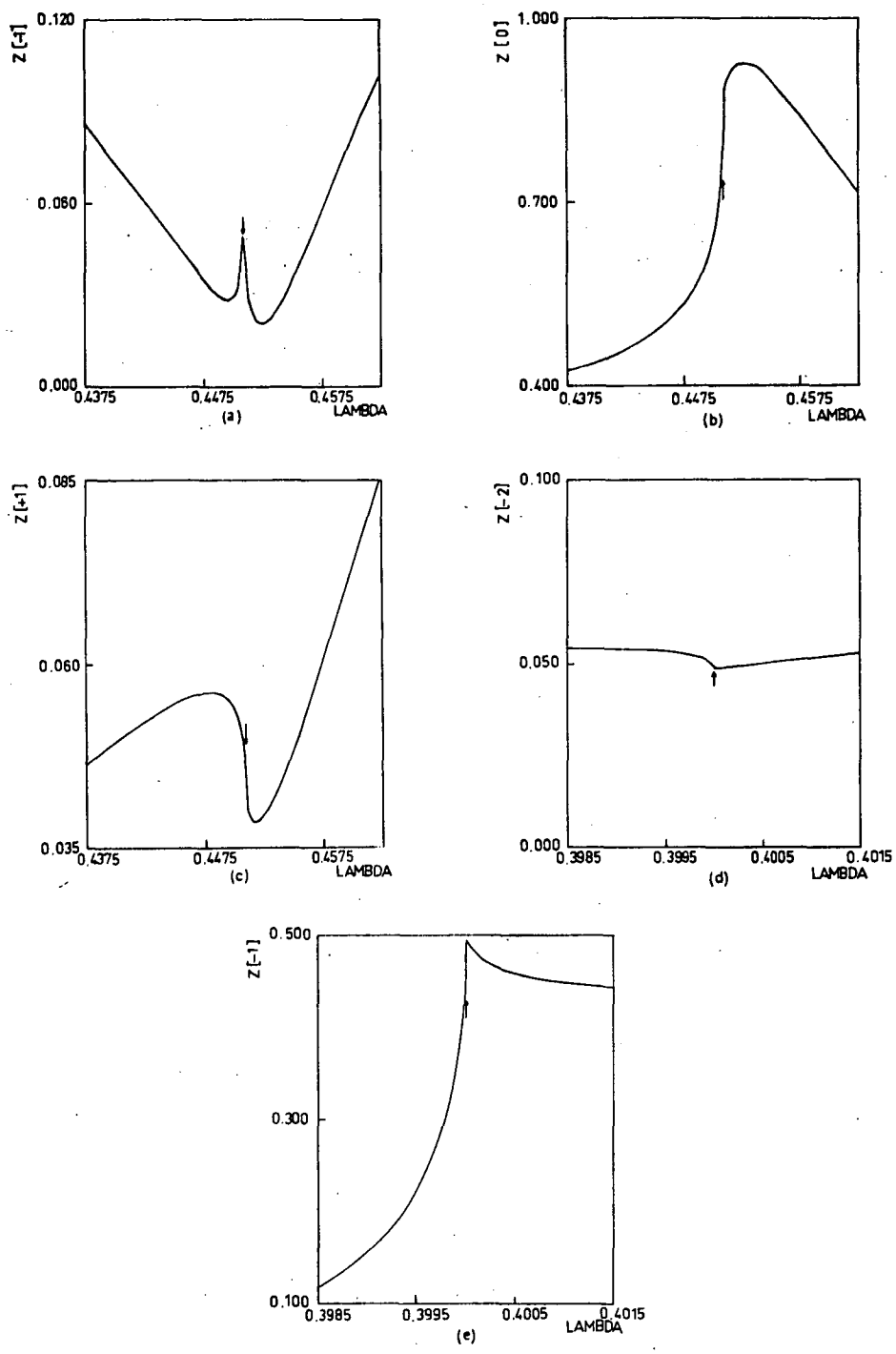


Figure 3.3. The effect of changing the order on the anomaly shape.

(a) for order -1]
(b) for order 0] correspond to a grating of profile depth 0.30 μm and period 1.00 μm, used at an angular deviation of 15 degrees in the order -1. The order -3 is passing off.
(c) for order +1]

(d) for order -2]
(e) for order -1] refer to the behaviour of the same grating at the simultaneous passing off of orders -3 and +2.

deviation of $+1.5^{o}$, and vice versa. This is what is meant by saying that the -3 and +2 anomalies exchange identities at an angular deviation of 0^{o}. Such an exchange of identities has been observed experimentally by Stewart and Gallaway (3.5).

The exact correspondence between the efficiency curves for the order -1 associated with the two non-zero angular deviations in fact follows from the second consequence of the Reciprocity Theorem.

The exchange of identities between the -3 and +2 anomalies can also be seen in the zeroth order (figure (3.4)(c)).

3.3.4. Transference of anomalies

We consider next the behaviour of the anomalies in the order -2 (figure (3.4)(a)). The lower wavelength anomaly for an angular deviation of 1.5^{o} consists of a pronounced step with two rounded edges. The upper anomaly is a much smaller step in the reversed sense to the lower step. For the opposite angular deviation, the lower anomaly is a symmetrical notch, while the upper one is a step ending in an efficiency maximum.

Thus, for the order -2 we do not have an exchange of identities between the anomalies as the angular deviation passes through zero. Instead, the anomalies of this order are transferred to the order +1. The anomalies in order +1 for an angular deviation of 1.5^{o} correspond exactly in wavelength, order and shape to those in order -2 for the opposite angular deviation. Similarly, the anomalies in order -2 for an angular deviation of 1.5^{o} correspond to those in order +1 for an angular deviation of -1.5^{o}.

This phenomenon of the transference of anomalies between different orders has not been previously reported.

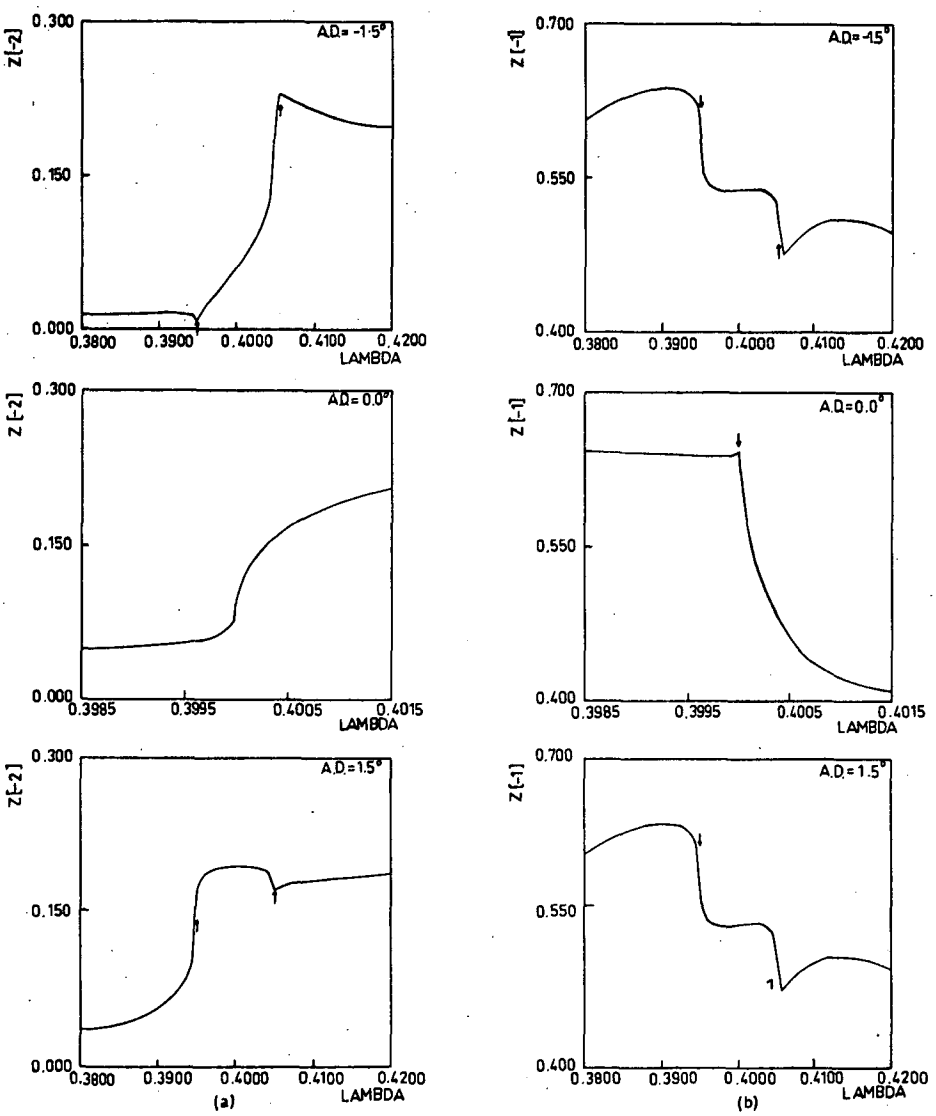

Z[-2]
0.300
0.150
0.000
A.O.= -1.5°
0.3800 0.3900 0.4000 0.4100 0.4200
LAMBDA
Z[-1]
0.700
0.550
0.400
A.O.= -1.5°
0.3800 0.3900 0.4000 0.4100 0.4200
LAMBDA
Z[-2]
0.300
0.150
0.000
A.O.= 0.0°
0.3985 0.3995 0.4005 0.4015
LAMBDA
Z[-1]
0.700
0.550
0.400
A.O.=0.0°
0.3985 0.3995 0.4005 0.4015
LAMBDA
Z[-2]
0.300
0.150
0.000
A.O.=1.5°
0.3800 0.3900 0.4000 0.4100 0.4200
(a)
LAMBDA
Z[-1]
0.700
0.550
0.400
A.O.=1.5°
0.3800 0.3900 0.4000 0.4100 0.4200
(b)
LAMBDA

Figure 3.4

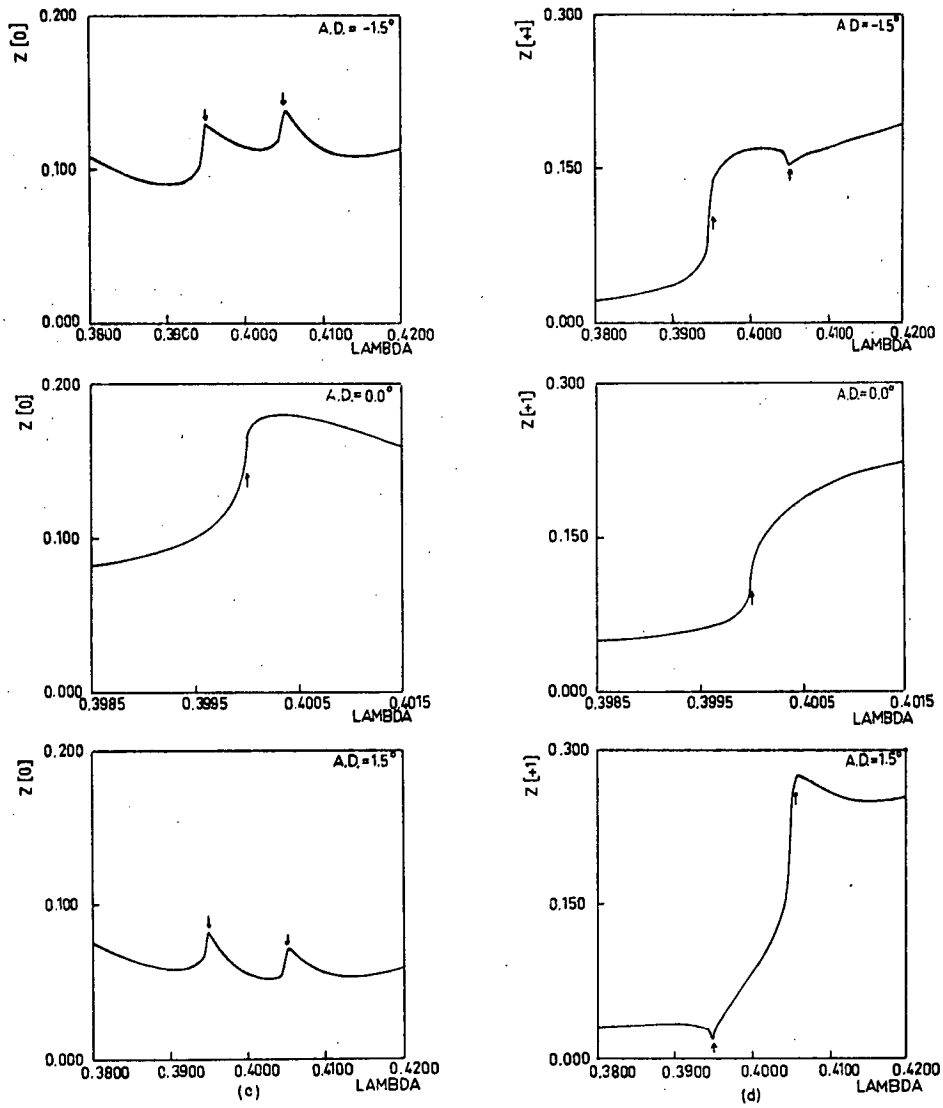


Figure 3.4. The exchange of identities between anomalies, and the transference of anomalies between two orders. The grating considered has period 1 μm, and profile depth 0.5 μm, and is used with the angular deviation (A.D.) constant in order -1. The three angular deviations represented here are -1.5 degrees, zero, and +1.5 degrees. (a) order -2; (b) order -1; (c) order 0; (d) order +1.

3.3.5 The location of the Rayleigh wavelength

In all the efficiency curves shown, the Rayleigh wavelength can be seen to be associated with the point of steepest gradient. This association has been well substantiated by experimental observations (3.5, 3.16, 3.17, 3.19, 3.20). It might well be expected on theoretical grounds also, since the kernel of the integral equation used in Pavageau and Bousquet's formulism involves the tangent of the angle of diffraction of the grazing order, which at passing-off is undergoing its most rapid rate of change.

3.4 TRIANGULAR PROFILE ANOMALIES

The triangular profile form is the most common one for diffraction gratings. Most of the experimental measurements of Wood anomalies have been made with gratings having this profile form.

The theoretical efficiency curves shown in this section all pertain to a grating having a period of 1.6667μ , groove angles of 10.4^{o} and 49.6^{o} and consequently a groove depth of 0.265μ.

3.4.1 The influence of the profile form on the anomalies

Figure (3.5) displays the Wood anomalies formed by this grating at the passing-off of the orders +1 (at 0.94462μ) and -2 (at 1.27147μ).

Figure (3.5)(a) shows an anomaly resembling in form the corresponding experimental anomaly (3.5), being a maximum having a very steep lower wavelength side on which the Rayleigh wavelength is located. Consequence (2) of the Reciprocity Theorem dictates that the anomaly be unaltered by the reversal of the sign of the angular deviation. This reversal was observed by Stewart and Gallaway to have a negligible effect on the anomaly.

If the grating groove angles are interchanged, the Wood anomaly changes markedly from a maximum to an edge of the form shown in figure (3.5)(b). The experimental anomaly was also affected by this interchange in a marked manner, although the off-blaze curve contains a minimum rather than the edge of figure (3.5)(b).

The on-blaze anomaly associated with the grazing order -2 is shown in figure (3.5)(c) to be a much stronger and sharper maximum than the +1 anomaly. This also is in agreement with the experimental findings.

The corresponding off-blaze anomaly is shown in figure (3.5)(d). This is stronger than the on-blaze anomaly. Stewart and Gallaway found this peak to be materially weaker in the off-blaze orientation. This must be attributed either to instrumental effects or the influence of the finite surface conductivity of the experimental grating. (Consequence (3) of the Reciprocity Theorem indicates that just after the Rayleigh wavelength, when only two orders are propagating, the efficiency in order -1 must be invariant with respect to the interchange of groove angles. This means that the peak efficiency values in the on-blaze and off-blaze orientations must be the same.)

3.4.2. Separation of Wood and resonance anomalies

The merging of the anomalies in order -1 associated with the passing-off of orders -2 and +1 is illustrated in figure (3.6). Figure (3.6)(a) shows the isolated form of each anomaly. When the angular deviation is halved (figure (3.6)(b)), the strength of the lower wavelength anomaly is diminished, while that of the other peak is unaffected. However, the Rayleigh wavelength now corresponds to the foot of the resonance peak, rather than a point on its lower

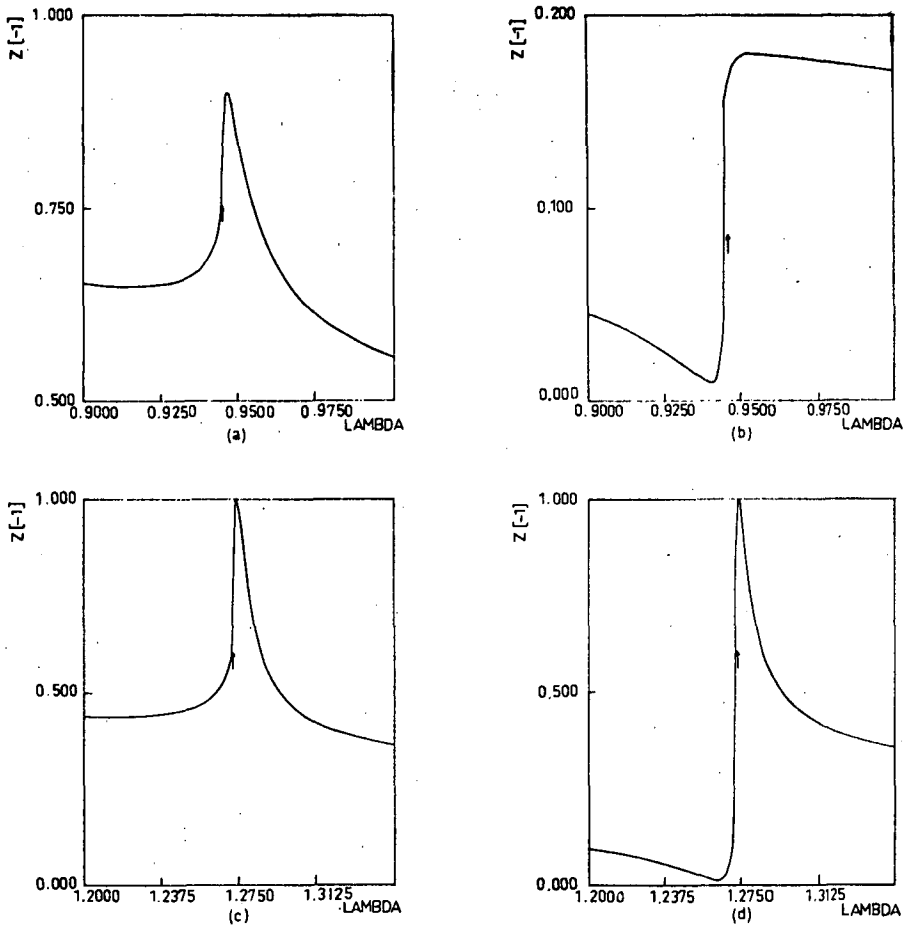


Figure 3.5. The effect of interchanging groove angles on triangular profile Wood anomalies. (These curves to be compared with those in figure 11 of (3.5).) (a) On blaze, order +1 passing off; (b) off blaze, order +1 passing off; (c) on blaze, order -2 passing off; (d) off blaze, order -2 passing off. The grating is used with the angular deviation in order -1 constant and equal to 18 degrees.

wavelength side. Figure (3.6)(c) shows that the process of weakening of the lower anomaly has continued, as has the separation of the upper Wood and resonance anomalies. Figure (3.6)(d), for an angular deviation of zero, has a near vertical edge, the combined Wood anomaly, separated from the combined resonance peak.

Thus, rather than the failure of Wood anomalies to combine as reported by Stewart and Gallaway (3.5), the process occurring here is one of combination of Wood and resonance anomalies, together with a separation of the two classes of anomaly.

Figure (3.6)(e) shows a corresponding experimental curve, for a silver-coated grating, taken from a paper by Strong (3.20). A similar curve was also obtained by Wood (3.19).

Hence, we have a new understanding of the double-band structure observed experimentally. The sharp edge of the dark band is the Wood anomaly, whose position is determined by the grating equation and must be independent of the nature of the grating surface. The bright band is a resonance peak, whose shape and strength are intimately linked with the equivalent surface impedance structure of the physical grating (3.4) and thus with the conductivity of the grating surface.

Figure (3.6)(f) shows the exchange of identities between the two anomalies reported by Stewart and Gallaway (3.5), and illustrates the second consequence of the Reciprocity Theorem.

3.4.3 Wavelength dependence of anomalies

The final property of the triangular profile anomalies which we will consider is the dependence of their form on the ratio of groove depth to Rayleigh wavelength.

Figure (3.7)(a) shows a double anomaly of different form to that of figure (3.6)(d) (for which the angle of incidence at the

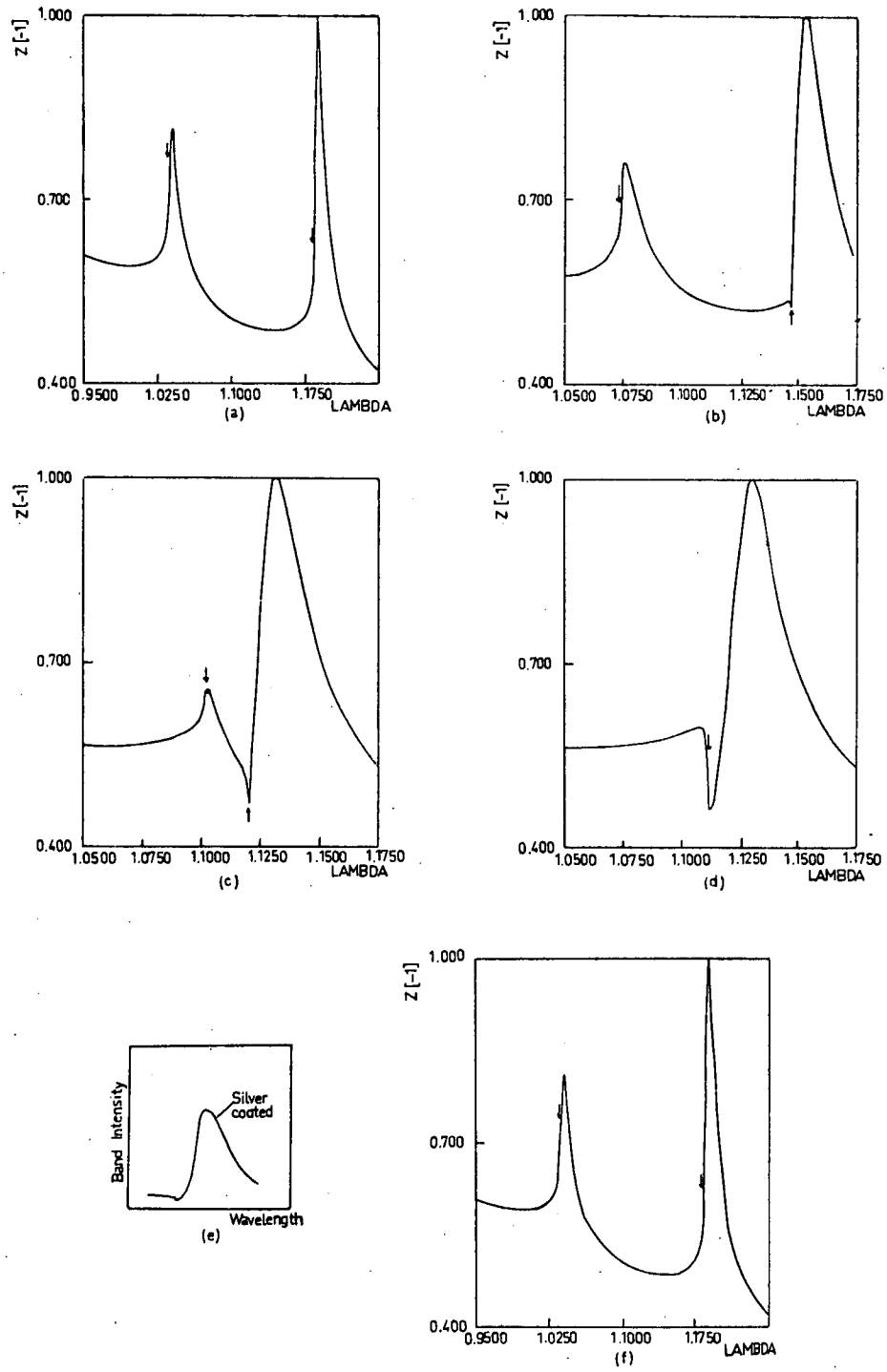


Figure 3.6. The merging properties of triangular profile Wood anomalies. The anomalies concerned are associated with the grazing orders -2 and +1. The angular deviation in the order -1 is kept constant for each curve. (a) A.D. = 8 degrees; (b) A.D. = 4 degrees; (c) A.D. = 1 degree; (d) A.D. = 0; (e) an experimental curve taken from figure 1 of (3.20); (f) A.D. = -8 degrees.

Rayleigh wavelength was 19.5^{o}, compared with a present value of 30^{o}). The Wood anomaly is now a narrow peak, while the resonance maximum is much less pronounced than it was previously.

For an angle of incidence of 40^{o} (figure (3.7)(b)), the single anomaly due to the grazing order +1 is of the form of an edge in the efficiency curve, rather than a resonance peak. It can be seen that the process of the weakening of the anomaly with decreasing Rayleigh wavelength has continued.

For an angle of incidence of 46^{o}, the anomaly due to the order +1 is so weak as to be barely perceptible. The dark anomaly which can be seen in figure (3.7)(c) is associated with the grazing order -6. It is of interest that for the zeroth order the opposite situation prevails, as figure (3.7)(d) shows. The anomaly due to the order +1 is reasonably strong, while that due to the order -6 is imperceptible. This is another illustration of the complex nature of the coupling between grazing and propagating orders.

Wood (3.19) found that the first-order bands which moved towards shorter wavelengths became less pronounced with increasing angle of incidence, and disappeared entirely at angles of incidence greater than 45^{o}.

The behaviour of the anomaly associated with the grazing-order -2 with increasing angle of incidence is in less satisfactory agreement with Wood's findings. The bands which moved further into the red were observed to become more pronounced (i.e. to show a greater change of intensity in passing from maximum to minimum) as the angle of incidence was increased.

Table (3.2) shows that the maximum change of efficiency (ΔZ) associated with the anomaly does not change greatly as the

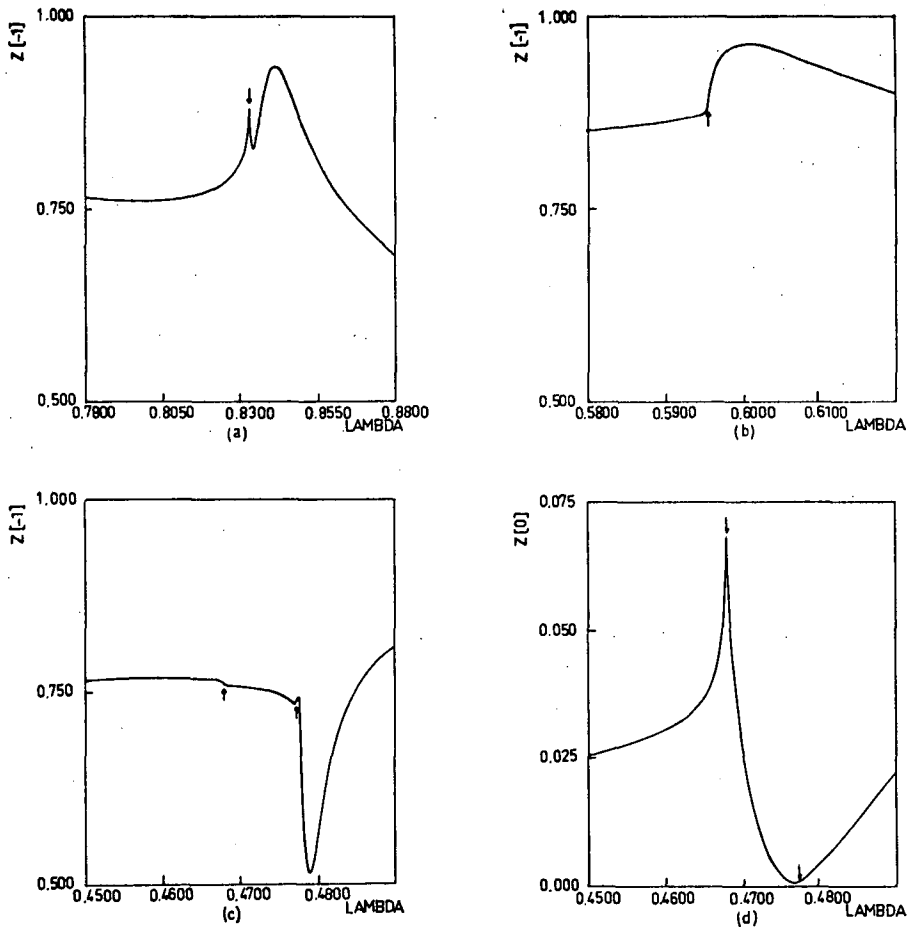


Figure 3.7. The effect of decreasing the Rayleigh wavelength on a Wood anomaly. (a) Angle of incidence 30 degrees, grazing orders +1 and -3, order -1, Rayleigh wavelength 0.83335 μm; (b) angle of incidence 40 degrees, grazing order +1, order -1,Rayleigh wavelength 0.59537 μm; (c) angle of incidence 46 degrees, grazing orders +1 (Rayleigh wavelength 0.46778 μm) and -6 (0.47760 μm), order -1. (d) As for (c), but the efficiency curve now corresponds to order 0.

Table (3.2)

Angle of incidence	$\lambda_R(\mu)$	ΔZ	$\Delta\lambda(\mu)$	R
32^0	1.2715	56 per cent	0.0094	135
40^0	1.3690	60 per cent	0.0065	211
46^0	1.4328	62 per cent	0.0063	228
60^0	1.5551	56 per cent	0.0060	259

Rayleigh wavelength (λ_R) increases. However, the theoretical anomalies do become more pronounced in the sense that their bandwidth ($\Delta\lambda$) decreases and their intrinsic resolving power ($R = \lambda_R/\Delta\lambda$) increases as the angle of incidence increases.

3.5. CONCLUSIONS

We have seen that a general theory of the diffraction by a grating can give good agreement with experimental observations of S-polarization diffraction anomalies. These 'anomalies' thus arise in a natural way from the well-established wave equation and boundary conditions on which the theoretical analysis is based, and require no special consideration.

The theoretical results have shown qualitative agreement with many experimental properties of anomalies. In some cases the agreement has been quantitatively accurate. This good quantitative agreement suggests that the theoretical process can adequately explain and predict the properties of these anomalies. However, further verification of this is needed.

There is a lack of observations of the anomalies of gratings having a known profile in the optical region. It is suggested that an investigation of the diffraction anomalies of a grating whose profile has been determined electron-microscopically should be undertaken, in order to provide an unequivocal test of the theory.

A good theoretical understanding of diffraction anomalies might well pave the way for the exploitation of their (in some cases) high intrinsic resolving power.

REFERENCES

(3.1) McPhedran, R.C., and Waterworth, M.D., 1972, Optica Acta, 19, 877.

(3.2) Wood, R.W., 1902, Lond.Edinb.Dubl.Phil.Mag., 4, 396.

(3.3) Twersky, V., 1956, I.R.E. Trans.Antennas Propag., 4, 330.

(3.4) Hessel, A., and Oliner, A.A., 1965, Appl.Optics, 4, 1275.

(3.5) Stewart, J.E., and Gallaway, W.S., 1962, Appl.Optics, 1, 421.

(3.6) Rayleigh, Lord, 1907, Proc.R.Soc. A, 79, 399.

(3.7) Fano, U., 1938, Annln Phys., 32, 393.

(3.8) Artmann, K., 1942, Z.Phys., 119, 529.

(3.9) Voigt, W., 1910, Nachr.Ges.Wiss.Göttingen, K, 1, 545.

(3.10) Fujiwara, S., and Iguchi, Y., 1968, J.opt.Soc.Am., 58, 361.

(3.11) Pavageau, J., and Bousquet, J., 1970, Optica Acta, 17, 469.

(3.12) Maystre, D., and Petit, R., 1970, Opt.Commun., 2, 309.

(3.13) Rayleigh, Lord, 1896, The Theory of Sound (London: Macmillan & Co. Ltd).

(3.14) Petit, R., 1967, Optica Acta, 14, 301.

(3.15) Brandes, R.G., and Curran, R.K., 1971, Appl.Optics, 10, 2101.

(3.16) Ingersoll, L.R., 1921, Phys.Rev., 17, 493.

(3.17) Palmer, C.H., 1952, J .opt.Soc.Am., 42, 269.

(3.18) Palmer, C.H., 1964, Symposium on Quasi-Optics (Brooklyn: Polytechnic Press).

(3.19) Wood, R.W., 1935, Phys.Rev., 48, 928.

(3.20) Strong, J., 1936, Phys.Rev., 49, 291.

CHAPTER 4

FURTHER PROPERTIES OF DIFFRACTION GRATING ANOMALIES

This chapter is based on a paper by the author and Dr. M.D. Waterworth (4.1), which has been accepted for publication by Optica Acta.

4.1. INTRODUCTION

The diffraction anomalies of deeply-ruled gratings have attracted little theoretical attention. The early observations by Wood (4.2) of anomalies in S polarized light were partially explained by Lord Rayleigh (4.3), using a perturbation treatment valid only for grooves shallow compared with the wavelength. Subsequent attempts (4.4, 4.5) to refine Rayleigh's theory failed to remove the restriction to gratings with shallow grooves.

This restriction was shown by Palmer (4.6) to have important consequences. His experimental studies demonstrated clearly the existence of diffraction anomalies in P polarized light. Such anomalies were not explicable by the above theories. Palmer attributed this failure to the circumstance that P anomalies are only formed by gratings with rulings deeper than those for which the plane wave expansion used by Lord Rayleigh, Fano and Artmann is valid.

It is only in recent years that theories of diffraction by gratings have become sufficiently sophisticated to deal with grooves deep compared with the wavelength. One such theory was used in Chapter 3 for a study of properties of S polarization Wood and resonance anomalies. In this chapter, the same theory is applied to the study of P polarization anomalies. Further attention is also

given to S anomalies.

After a discussion of a useful theoretical result, a comparison is made between our calculated values, measurements at millimetre wavelengths and values based on two recently-developed theoretical formulisms (4.7, 4.8). Some properties of P anomalies for gratings having both sinusoidal and triangular profiles are discussed. Finally, deep sinusoidal gratings are shown to furnish higher-order resonance anomalies for S polarized light. These anomalies are extremely pronounced for narrow ranges of groove depths, described by an equation first suggested by Hessel and Oliner (4.9).

4.2. THEORETICAL CONSIDERATIONS

The theoretical method used here for the calculation of grating efficiencies is based on the integral equation formulation of Pavageau and Bousquet (4.10), and has been described in Chapter 2 and Appendix I. All the results based on this formulation and quoted below are accurate to better than ±1%, except where stated.

In considering graphs in which diffraction grating efficiency is plotted as a function of angle of incidence (θ) for a constant wavelength (λ), a corollary of the Reciprocity Theorem is often of use. Let E(n) denote the diffraction efficiency in order n, and let θ_0 be that angle of incidence for which the angle of diffraction in order n is $\theta_n = \theta_0$. Then E(n) is a symmetrical function of θ about θ_0. for small excursions from θ_0 (if $n \neq 0$). Note that E(0) is exactly symmetrical about $\theta_0 = 0$.

This symmetry property has been previously stated by Palmer and Le Brun (4.8). However, they did not make clear that the symmetry was confined to a small region about the angle θ_0.

The usefulness of this property is shown in Figure (4.1), in which lines of symmetry are indicated for five orders (2, 1, 0, -1, -2). These S polarization efficiency curves have been calculated using the data from Figure (11) of reference (4.7). The two sets of curves are generally in agreement, except in the vicinity of Wood anomalies, where in some cases serious discrepancies occur. The most important discrepancy occurs near the angle of incidence (-26.75^{o}) where orders +2 and -1 pass off. The calculations of Jovicevic and Sesnic indicate here an efficiency peak reaching up to 100% in order 0. However, there exists no corresponding efficiency peak near $\theta = 26.75^{o}$. Evidently, Figure 11 contains localized features not in agreement with the Reciprocity Theorem, which has been verified both theoretically and experimentally (4.11). No such features are present in Figure (4.1).

The existence of these localized discrepancies between the two sets of theoretical results can be attributed to a difference in the angular resolutions of the two calculations. Jovicevic and Sesnic indicate that the angular resolution for their S polarization efficiency curves was between 2^{o} and 2.5^{o}. The angular resolution of the calculations reported here is better than 0.1^{o} in the region of Wood anomalies.

Confinement of the symmetry of efficiency curves of non-zero orders to small angular excursions can be seen in Figure (4.2). The P polarization efficiency curves for order +2 are locally symmetrical about $\theta = -30^{o}$. However, this order passes off for an angle of incidence of 0^{o}, and has an efficiency of the order of 10% for $\theta = -60^{o}$. The symmetry in this order is thus limited by the diffraction grating

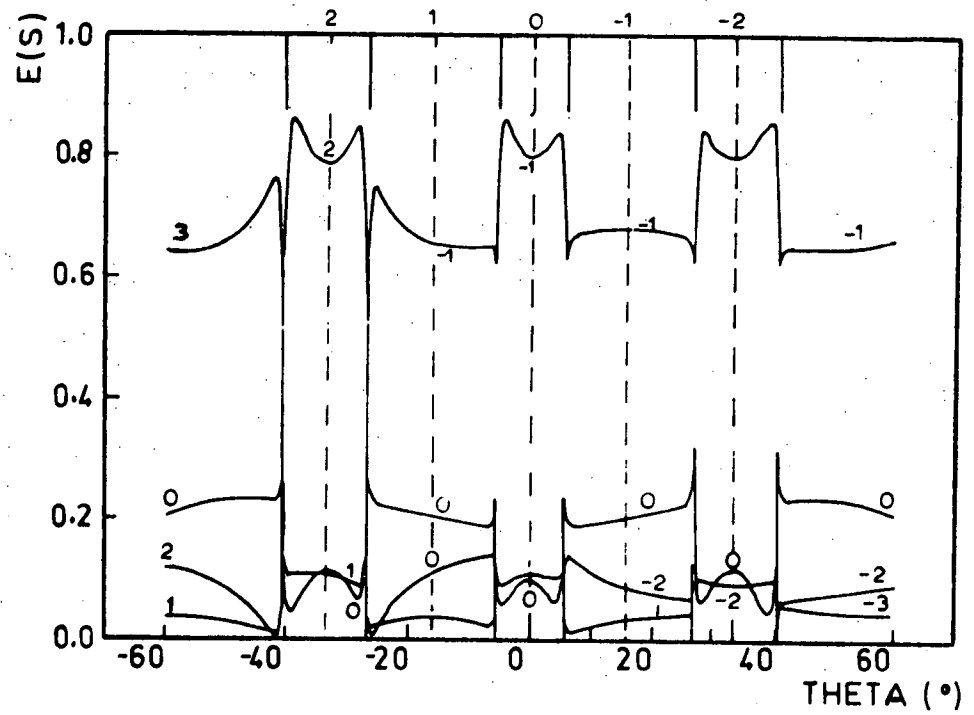


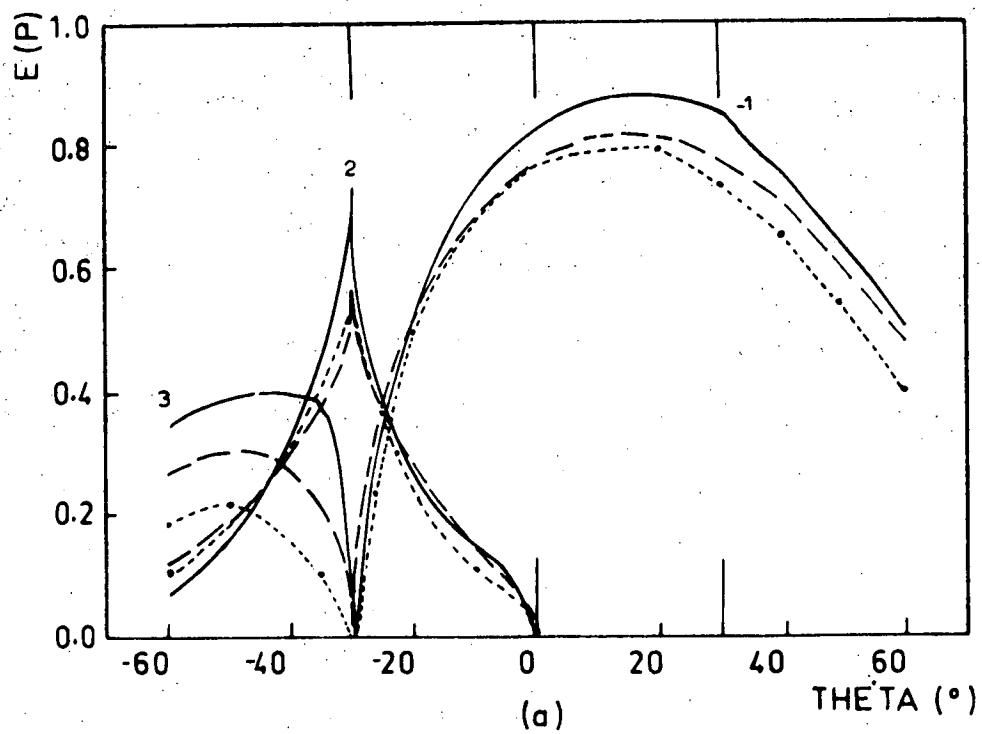
Figure (4.1). S polarization efficiency curves for a triangular profile grating with groove angles 10^{0} and 80^{0} and period d, used with light of normalized wavelength $\lambda/d = 0.55$. The dashed lines represent lines of symmetry for the various orders; the numeral at the top of each dashed line indicates the order concerned. As in the following figures, Rayleigh angles (where orders pass off or arise) are indicated by the longer lines perpendicular to the theta axis.

equation to no more than an angular range of 30^{o} about $\theta = -30^{o}$.

4.3. COMPARISONS OF THREE THEORIES WITH EXPERIMENT

Figure (4.2) shows theoretical efficiency curves based on two different formulisms (4.7, 4.10), as well as experimental measurements made at millimetre wavelengths (4.12). All three sets of curves correspond to the same triangular grating profile, and to P polarized radiation. The experimental curves and those based on the series-expansion formulism of Jovicevic and Sesnic are taken from Figure 3 of reference (4.7).

For order 3, the experimental curve reaches a peak efficiency of just over 20%. The curve of Jovicevic and Sesnic attains a maximum of 40%, whereas our curve has a maximum value of 30%. For order 2, all three curves show a P polarization Wood anomaly in the form of a cusp-like maximum. The reason for the anomaly taking this particular form will be discussed more fully below. The peak experimental efficiency is 50%; compare this with the peak value of Jovicevic and Sesnic of about 73%. Our maximum value is 57%. The angular resolution of the experimental measurements corresponded to a half-angle of 2^{o}. This limited resolution is sufficient to account for the discrepancy of 7% in peak efficiency values. For order 1, both theoretical curves have two peaks which are sharper than those of the experimental curve. The experimental points, however, appear to be insufficiently closely spaced around the peaks to define their shape accurately. Our points lie in the main between the two other curves. The deviation between the three curves as the angle of diffraction approaches 90^{o} corresponds to the experimental angular bandwidth. Again for order 0 our curve lies between the other two


E (P)
1.0
0.8
0.6
0.4
0.2
0.0
-60
-40
-20
0
20
40
60
THETA (°)
1
2
3
(a)

curves. This is the case for order -1, for angles of incidence larger than -20^{o}. The peak experimental efficiency in this blazed order is 79%. Our peak value is 82%. The other theoretical value is 88%. For angles of incidence between -20^{o} and -30^{o}, the discrepancy between the three curves is equivalent to an angular error of no more than 3^{o}.

Thus, on the whole, the results of the integral equation formulation are in better agreement with the experimental values than are the results of Jovicevic and Sesnic. Where the agreement is worse, the discrepancy is only slightly larger than the experimental error. The worst discrepancy between our theory and experiment is about 12% (for order -3). However, for the predominant P anomaly of Figure (4.2), the agreement between our theory and experiment is accurate to within the errors of the latter.

Figure (4.3) shows theoretical efficiency values based on the integral equation formulism, as well as theoretical values and measurements made at millimetre wavelengths, taken from Figures 4 - 7 of reference (4.8). The theoretical values of Palmer and Le Brun are obtained using the standard Kirchhoff integral, modified to take into account polarization effects. They have used one arbitrary scaling factor for each polarization to give best agreement between theory and experiment. No such scale factors have been used for the integral equation values.

The three curves are in generally good agreement for the zeroth order and P polarized radiation. However, Palmer and Le Brun's theoretical curve deviates markedly from the other two for angles of incidence larger than 25^{o}, near the edge of the region of validity of their theory. The other theoretical curve is in good agreement with

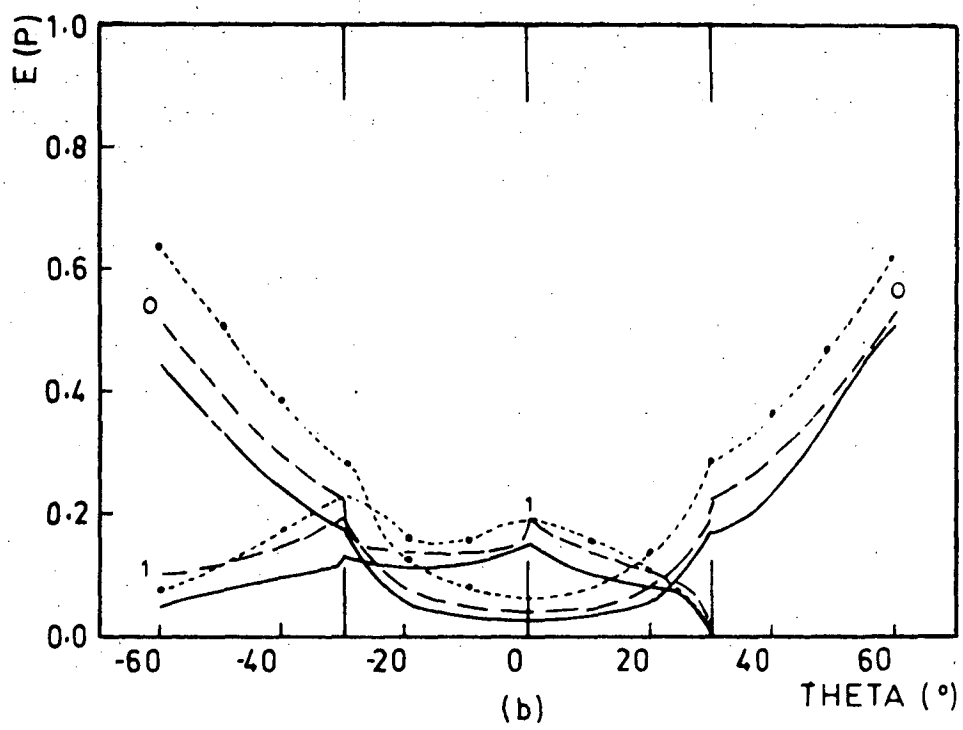
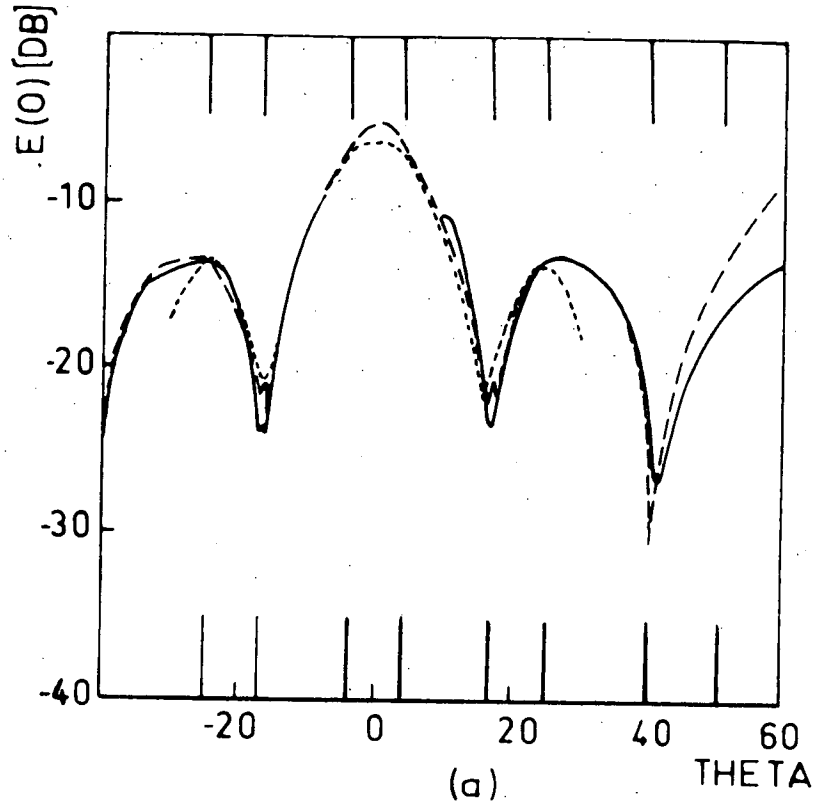
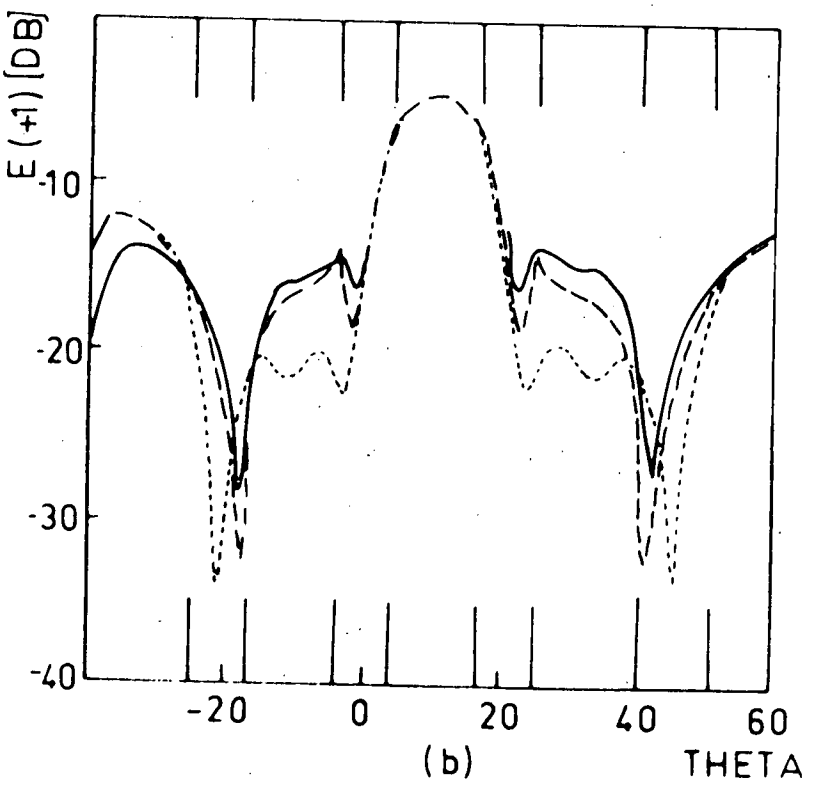
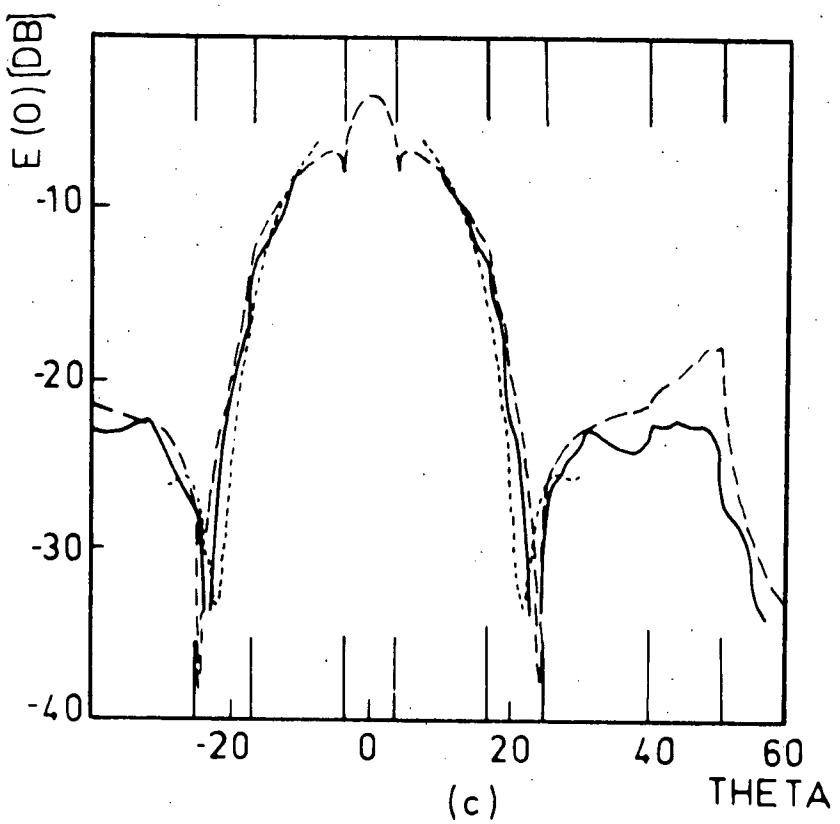


Figure (4.2). P polarization efficiency curves for a triangular profile grating with groove angles 15^{o} and 75^{o}, used with light of normalized wavelength $\lambda/d = 0.50$. The solid curves correspond to values calculated by Jovicevic and Sesnic, while the fine-dashed lines join the experimental points of (4.12). The longer-dashed lines outline the curves corresponding to integral equation values. (a) Orders 3, 2 and -1; (b) Orders 0 and +1.


E(0) [DB]
-10
-20
-30
-40
-20
0
20
40
60
THETA
(a)


E(+1) [DB]
-10
-20
-30
-40
-20
0
20
40
60
THETA
(b)


E(O)[DB]
-10
-20
-30
-40
-20
0
20
40
60
THETA
(c)

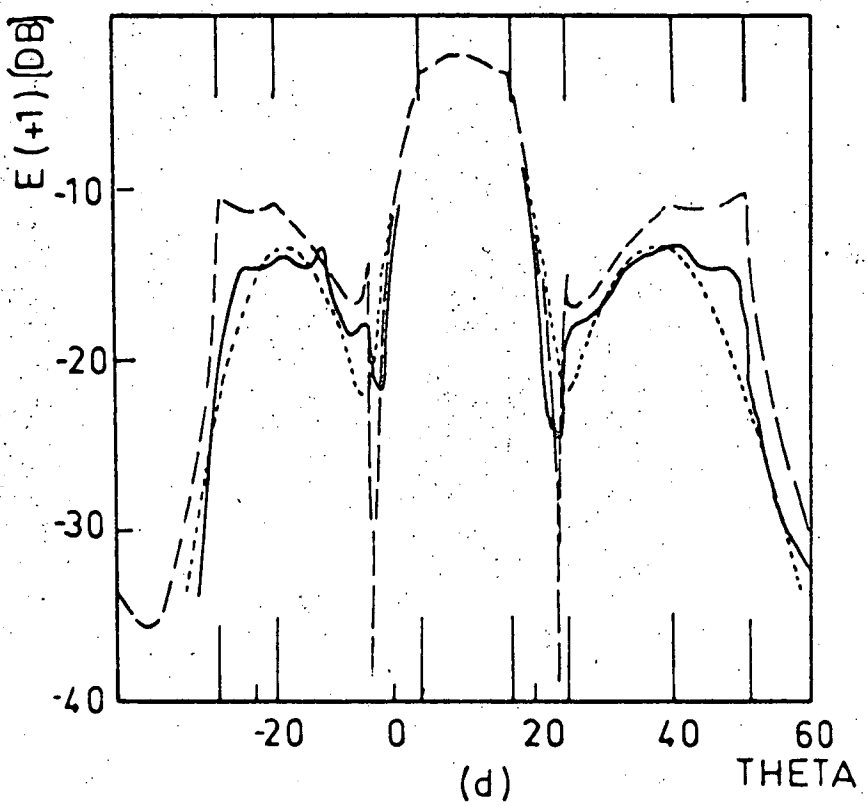


Figure (4.3). Efficiency curves for a triangular profile grating with groove angles 30^{o} and 60^{o}, used with light of normalized wavelength $\lambda/d = 0.355$. The fine dashed lines correspond to values calculated by Palmer and Le Brun, while the solid curves indicate their experimental measurements. The longer-dashed lines outline the curves corresponding to integral equation values. The efficiency scales are in decibels. (a) Order 0; P polarization. (b) Order +1; P polarization. (c) Order 0; S polarization. (d) Order +1; S polarization.

the experimental curve for angles of incidence up to 40^{o}. Above 40^{o}, the discrepancies between the two grow in magnitude, although the two curves still have similar shapes. Note the successful prediction by the integral equation theory of an anomalous maximum in the centre of the wide dip near $\theta_i = -17^{o}$. The experimental measurements do not indicate the presence of another such maximum near $\theta_i = 17^{o}$, as would be expected from the Reciprocity Theorem. This must therefore be attributed either to experimental error or to the effects of the finite conductivity of the experimental grating's surface.

For order +1 and P polarization, the theoretical curve of Palmer and Le Brun deviates considerably from the other two curves. Both theoretical curves indicate deeper anomalous minima corresponding to the passing off of orders 2 and -1 than can be seen on the experimental curve. However, the angular resolution of the millimetre wave telescope used by Palmer and Le Brun was about 0.5^{o}. Thus their measured values were in fact efficiencies averaged over a 0.5^{o} range of diffraction angles. The effect of such an averaging is to make anomalous minima more shallow, and also to smooth out spikes such as those predicted by the integral equation formulation. This method correctly determines the positions of the four anomalous minima, whereas angular discrepancies can be seen between the experimental and theoretical curves of Palmer and Le Brun.

For order zero and S polarization, the two theoretical curves agree reasonably well with the experimental curve in the region between the anomalous minima. The effects of the limited experimental angular resolution are sufficient to account for the discrepancies in the strength of the minima. The integral equation formulation

evidently has a larger angular range of validity than the modified Kirchhoff method. For the former, the discrepancies between theory and experiment become significant for angles of incidence larger in magnitude than 30^{o}.

For order +1 and S polarization, the theoretical curve of Palmer and Le Brun has a general form more in keeping with the experimental results than has the other theoretical curve. However, the first curve has a smooth form which does not show most of the anomalous features present on the other two curves. The integral equation formulism successfully predicts the formation of an upward spike on the edge of the deep anomalous minima. Again, the averaging effect of the limited experimental angular resolution is sufficient to reconcile these two curves in the neighbourhood of these pronounced S polarization anomalies. Discrepancies between the integral equation curve and the experimental curve tend to increase for angles of incidence larger than 30^{o}; a similar effect occurs when the angle of diffraction increases beyond 35^{o}. As pointed out by Palmer and Le Brun (4.8), the effects of the finite conductivity of the grating surface would be expected to be most significant for S polarized radiation at high angles of incidence and diffraction.

Thus, for P polarization the integral equation formulism is seen to yield results in better agreement with experiment than those from the modified Kirchhoff theory. For S polarization, the two theories are of comparable accuracy, although the former gives the better predictions with regard to the shape of anomalous features. The region of validity of the integral equation formulation is larger than that of the Kirchhoff theory. Both theories seem to lose some accuracy at high angles of incidence and diffraction. This may be a consequence

of the finite conductivity of the grating surface.

The P polarization anomalies of Figures (4.2) and (4.3) are much broader than the S anomalies of Figures (4.1) and (4.3). This is a general feature of P anomalies, which has been observed experimentally (4.13). As a consequence, greater angular accuracy is required both of theoretical calculations and of experimental observations for S polarization than for P polarization.

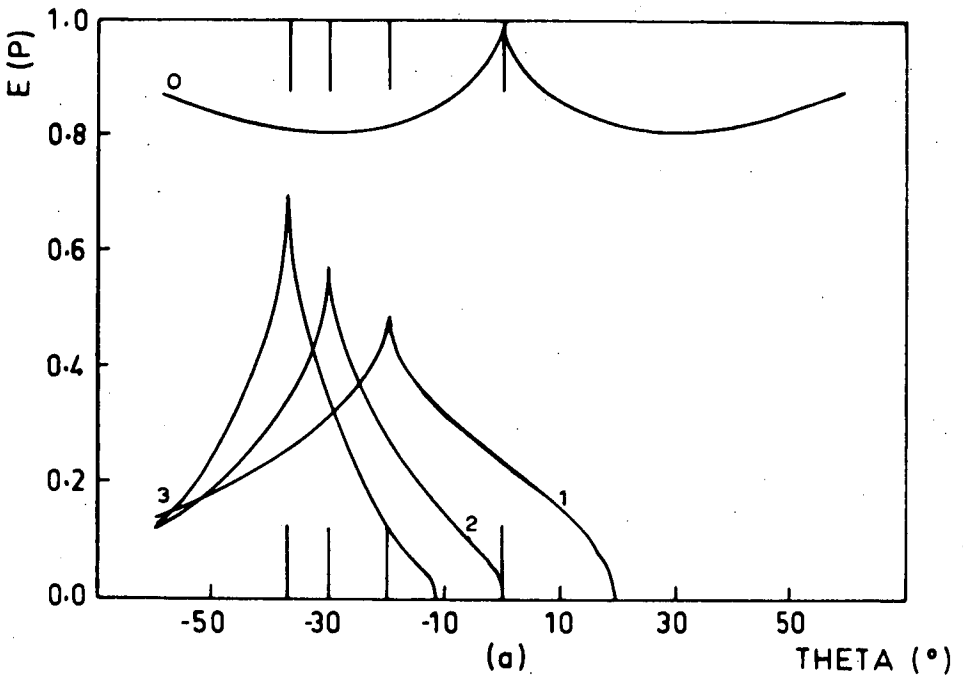
4.4. PROPERTIES OF P ANOMALIES

The most pronounced P anomaly of Figure (4.2) was the cusp-like maximum in order 2. The reason for the anomaly taking this particular form is clear. As mentioned above, the Reciprocity Theorem ensures in this case that E(2) is an approximately symmetrical function about $\theta = -30^{o}$. Since the blazed order -1 passes off at this angle, together with order 3, the efficiency of order 2 would be expected to curve upwards around $\theta = -30^{o}$, in order to preserve the energy balance. Since there is an association between Rayleigh angle and steep gradients of the efficiency curve for P anomalies just as for S anomalies (4.6, also Chapter 3), the maximum in order 2 must assume the symmetrical, pointed form of Figure (4.2).

Thus, we would expect a strong P anomaly in the form of a cusp-like maximum whenever the blazed order -1 passes off simultaneously with a positive order (m_+). From reference (4.14), this will occur for gratings of unit period when

$$\lambda = 2/(m_+ + 1) \quad .$$

Figure (4.4) shows that P anomalies of the predicted form do occur for $m_+ = 1,2,3,4$. As the wavelength decreases, the anomalies


E (P)
1.0
0.8
0.6
0.4
0.2
0.0
0
1
2
3
-50
-30
-10
10
30
50
(a)
THETA (°)

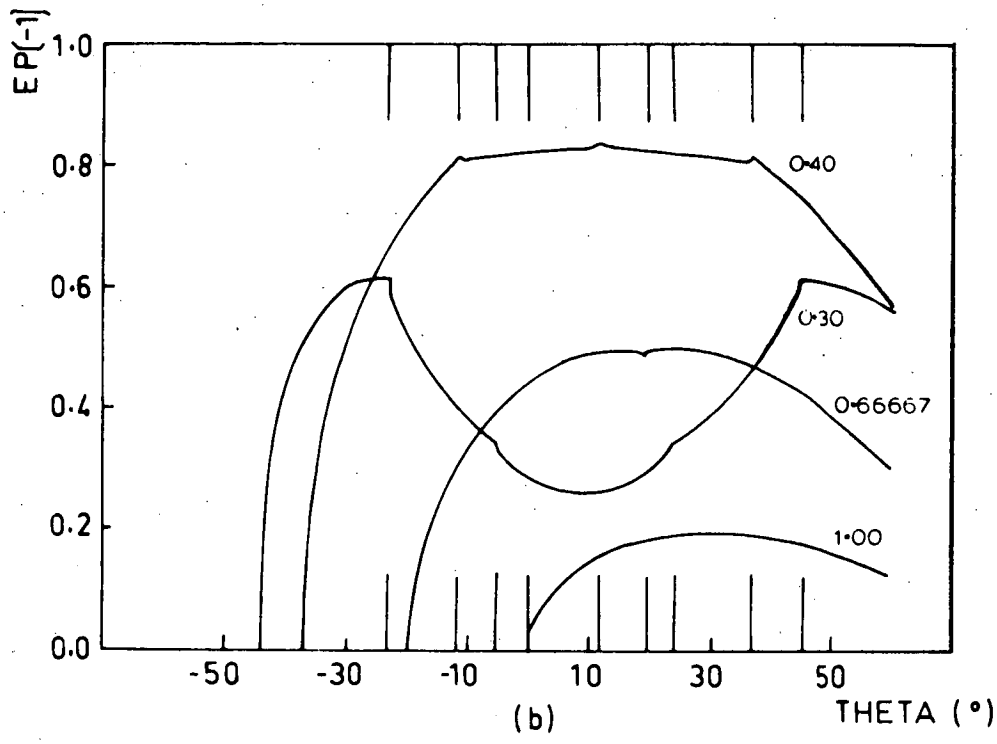
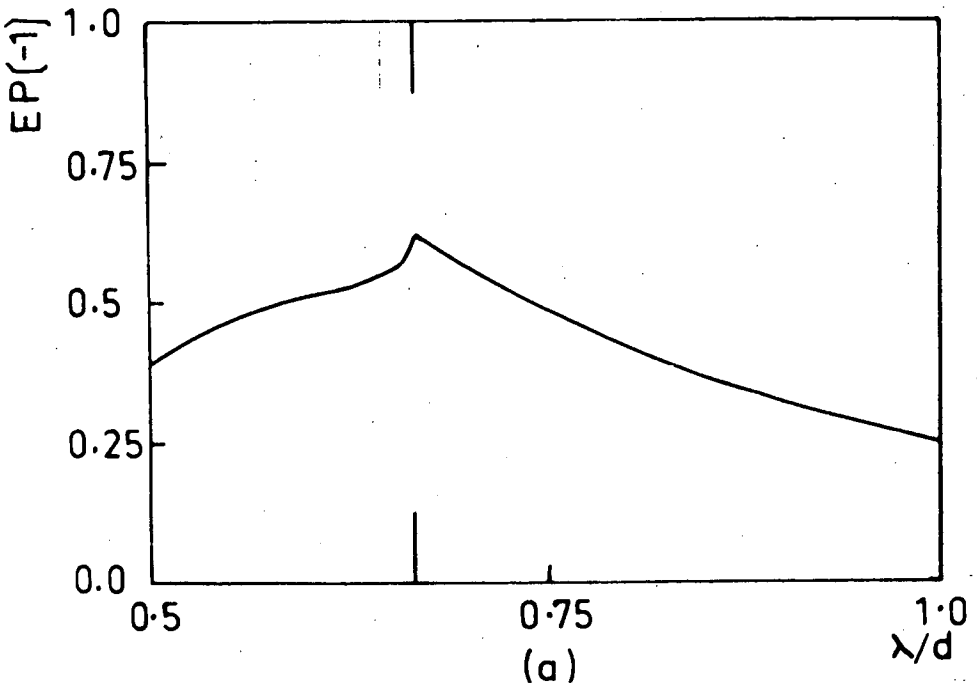


Figure (4.4). P polarization efficiency curves for a triangular profile grating with groove angles 15^{o} and 75^{o}. (a) The development of an anomalous cusp with decreasing wavelength. Orders 0, 1, 2 and 3 correspond respectively to normalized wavelengths 1.00, 0.66667, 0.50 and 0.40. (b) The development of anomalies in the blazed order -1 with decreasing wavelength. The curves shown are for normalized wavelengths of 1.00, 0.66667, 0.40 and 0.30.

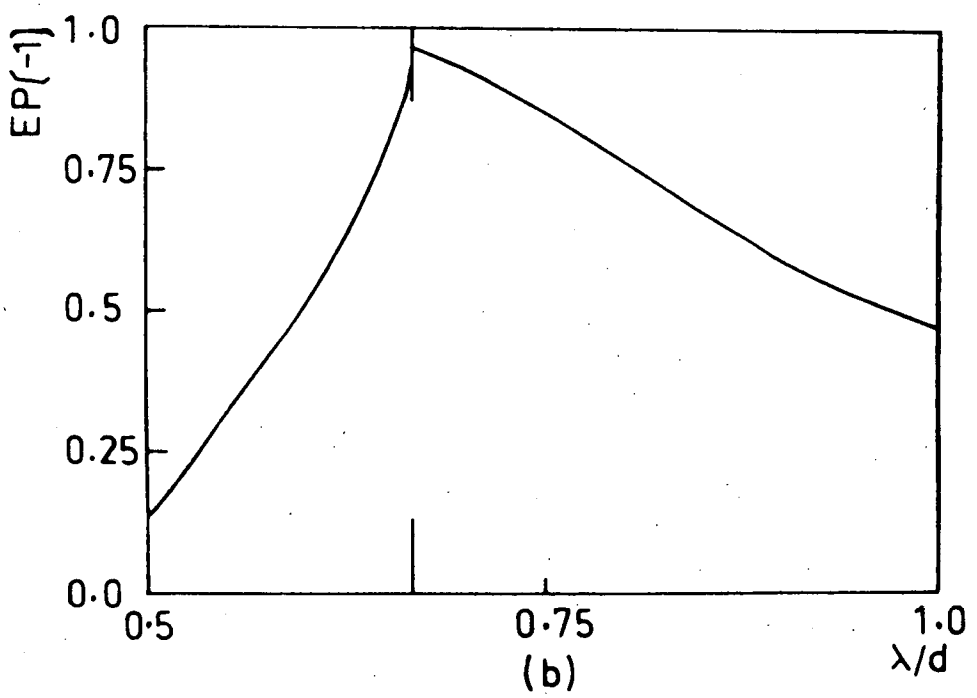
become stronger and sharper. This is in agreement with the experimental observations of C.H. Palmer (4.6). However, significant P anomalies are formed by this grating for wavelengths up to four times its groove depth of 0.25 μm. Palmer found that P anomalies tended to be formed only for wavelengths smaller than or comparable to the groove depth. Figure (4.4) shows that the P anomalies in the blazed order -1 do in fact only become strong when the wavelength decreases to 0.30 μm, close to the groove depth in magnitude. The strengthening of this anomaly occurs simultaneously with a strengthening of the neighbouring order -2. For this grating, the strongest P anomalies tend to be formed in the order adjacent to a strong passing-off order.

Consider next the variation of the strength and form of the P anomaly caused by the passing off of orders -2 and +1, for a grating having a sinusoidal profile (of depth D = 2A). In a study of the blaze properties of the sinusoidal profile, this anomaly was seen to play a key role in the attainment of good spectral performance in unpolarized light from sinusoidal gratings (refer to Chapter 6).

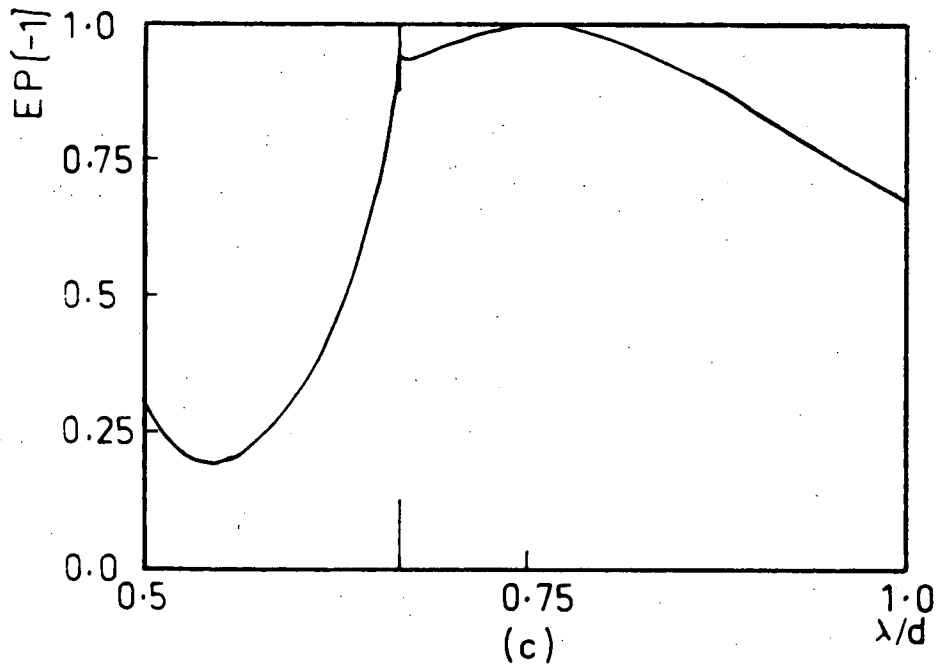
This anomaly is barely perceptible for A = 0.05, but is clearly evident for A = 0.10 (Figure (4.5)). For this profile depth, the Rayleigh wavelength is associated with a branch-point of the efficiency curve, and also with a local maximum of the gradient of the efficiency curve. Such branch-point anomalies have been considered previously for S polarization in Chapter 3. By A = 0.15, the P anomaly has strengthened further, and has become the dominant influence on the form of the P polarization efficiency curve. For A = 0.20, the anomaly has two peaks. One of these terminates the steep efficiency curve segment around the Rayleigh wavelength. The other is more rounded, and corresponds to the P polarization resonance peaks in the

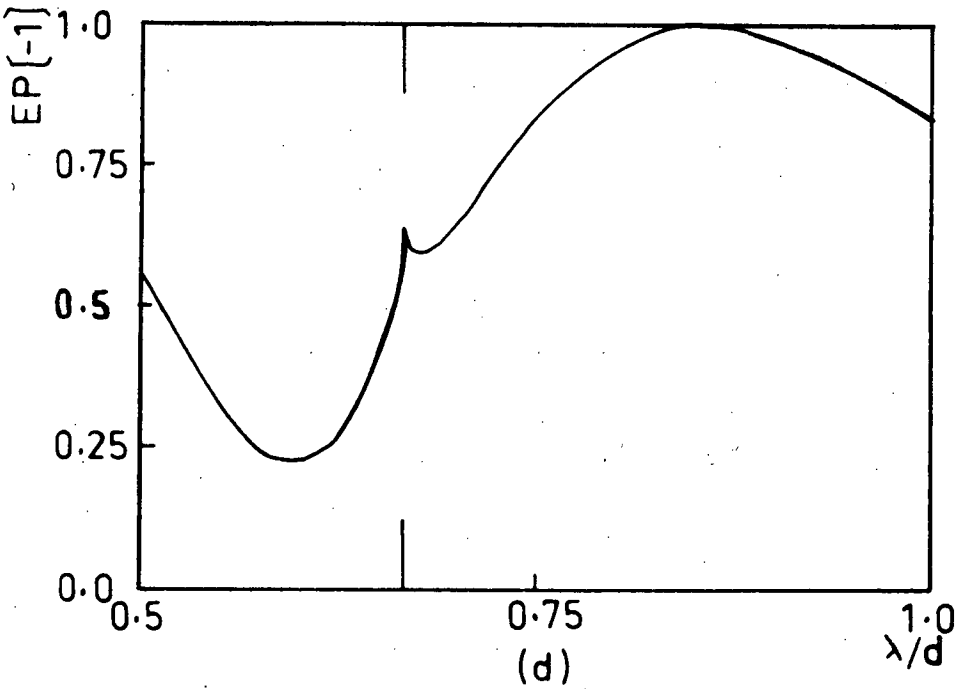


(a)



(b)


EP(-1)
1·0
0·75
0·5
0·25
0·0
0·5
0·75
1·0
λ/d
(c)


EP(-1)
1·0
0·75
0·5
0·25
0·0
0·5
0·75
1·0
λ/d
(d)

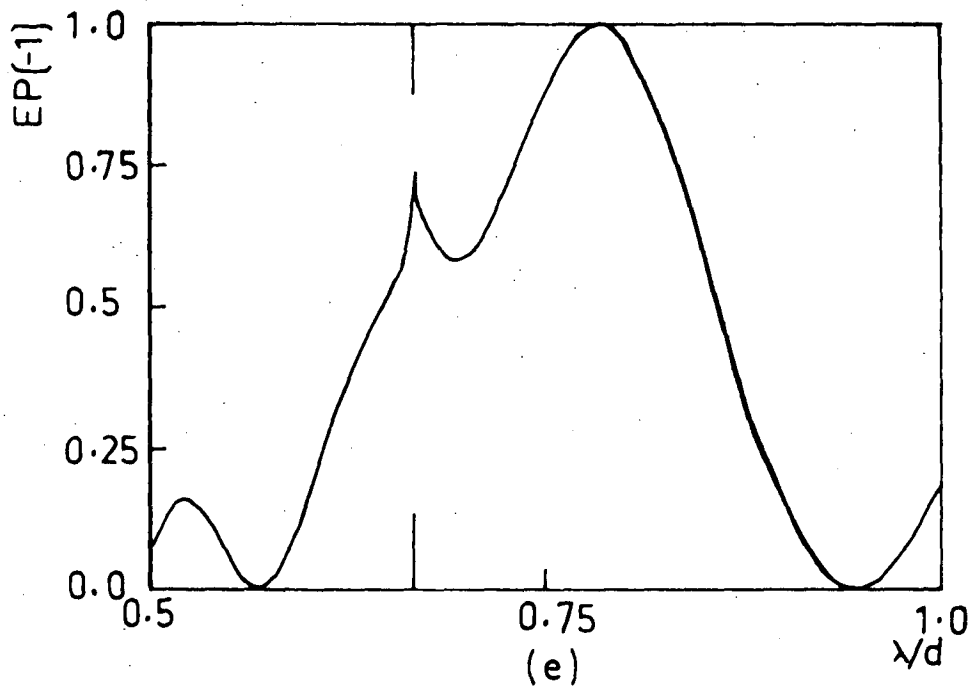


Figure (4.5). P polarization efficiency curves for sinusoidal gratings used in a Littrow mount in the order -1. In these graphs, and those of figure 9, the Rayleigh wavelength is indicated by the longer lines perpendicular to the axis of λ/d. The normalized profile amplitudes (A/d) are, respectively, (a) 0.10, (b) 0.15, (c) 0.20, (d) 0.25 and (e) 0.65.

evanescent orders -2 and +1. Such resonance anomalies have been discussed in detail for S polarization in Chapter 3 and reference (4.9), but P polarization resonances have not previously been considered. By A = 0.25, the separation of the Wood and resonance anomalies has progressed to such a stage that the predominance of the latter is clear. As the groove depth increases further, the Wood anomaly weakens, while the P resonance maximum remains largely unchanged in shape and strength, but moves towards longer wavelengths. The Wood anomaly passes through a minimum strength, and after A = 0.40 rapidly becomes more prominent. The final curve of Figure (4.5), corresponding to A = 0.65, shows that a second resonance peak has begun to detach itself from the Wood anomaly.

Hence, the P polarization Wood and resonance anomalies behave in an oscillatory fashion as the depth of the sinusoidal profile is varied. The Wood anomalies grow and then decline in strength as the depth varies through certain ranges. The behaviour of the resonance anomalies evidently plays a key role in the determination of this property.

This oscillatory behaviour can be understood in terms of an equation given by Hessel and Oliner (4.9). According to them, in order for strong P anomalies to be formed, the grating groove depth D must be such that the equivalent surface impedance structure is capacitive. If λ_g denotes the wavelength of the P polarization guided wave which can be supported by the surface, then the requirement of a capacitive structure is expressed by the equation

$$\frac{n\lambda_g}{2} > D > (2n-1)\frac{\lambda_g}{4} \quad , \quad n = 1, 2, \ldots$$

Thus P anomalies would be expected to fluctuate in strength as the groove depth moves in and out of the regions in which the above inequality

holds.

It should be noted that it is the strength of the P Wood anomaly which waxes and wanes with the profile depth. Once the resonance maximum in order -1 has reached full strength, it retains that strength.

Figure (4.6) shows the form of the variation of the maximum intensities of the P resonances in the evanescent orders -2 and +1 with grating groove amplitude. Both maximum intensities increase steadily with groove depth, and to a good approximation vary according to the exponential law

$$I_{max} \propto 10^{6.32(A/d)}$$

where A/d is the groove amplitude divided by the groove period. I_{max} thus increases by a factor of 2.07 for every increase of A/d by 0.05.

The P resonance peak in order -1 is formed by the joint action of the two resonances in the evanescent orders. The wavelength of peak efficiency always lies between the two wavelengths of peak intensity, and all three wavelengths increase in similar smooth fashions with groove amplitude.

According to Palmer (4.8), the polarization of the light diffracted into the grazing orders is a good indicator as to the possibility of formation of P anomalies. If P anomalies are to exist, a significant fraction of the tangentially diffracted light must be P polarized. Figure (4.7) shows the variation of the polarization of the two grazing orders -2 and +1 with the normalized amplitude of the sinusoidal profile. For both orders, ES/EP initially decreases steadily with A/d. In the first region of strong P anomalies in order -1 (A/d $\simeq$ 0.15) about 10% of the tangentially diffracted light is

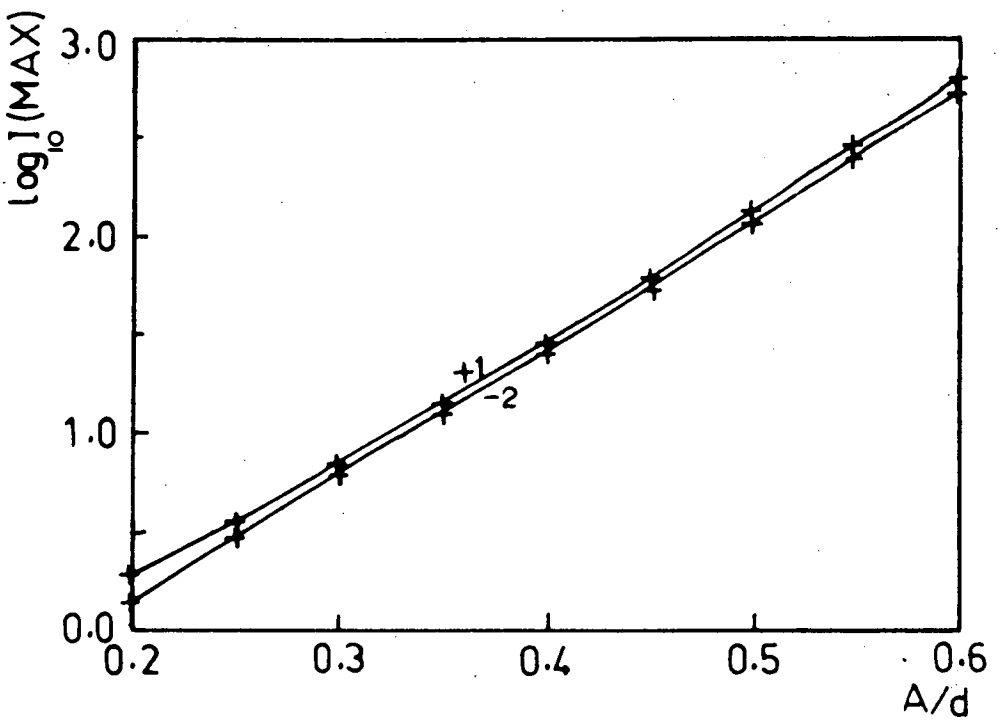


Figure (4.6). The variations of the logarithms of the maximum intensities of the P polarization resonance peaks in the evanescent orders -2 and +1 are shown, as functions of the normalized amplitude of the sinusoidal grating profile.

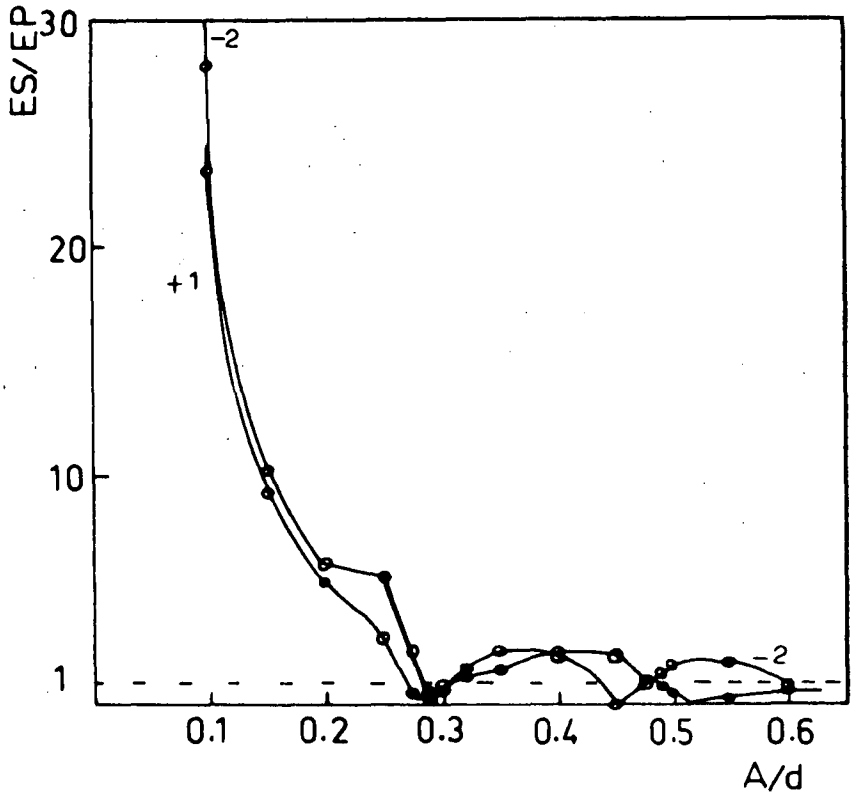


Figure (4.7). The variations of the polarizations of the grazing orders -2 and +1 are shown, as functions of the normalized amplitude of the sinusoidal grating profile. ES and EP denote, respectively, the S and P polarization efficiencies corresponding to a normalized wavelength of 0.666, and an angle of diffraction of magnitude 87.4^{o}.

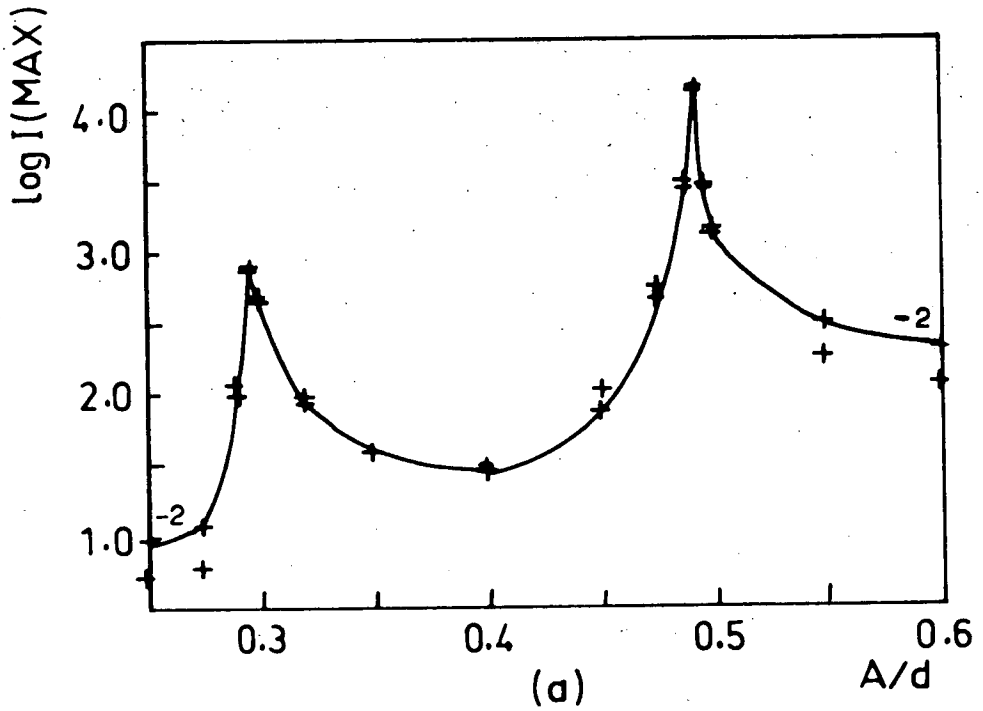
P polarized. In the second region of strong P anomalies (A/d ≃ 0.55) the average value of ES/EP for the two grazing orders is close to one.

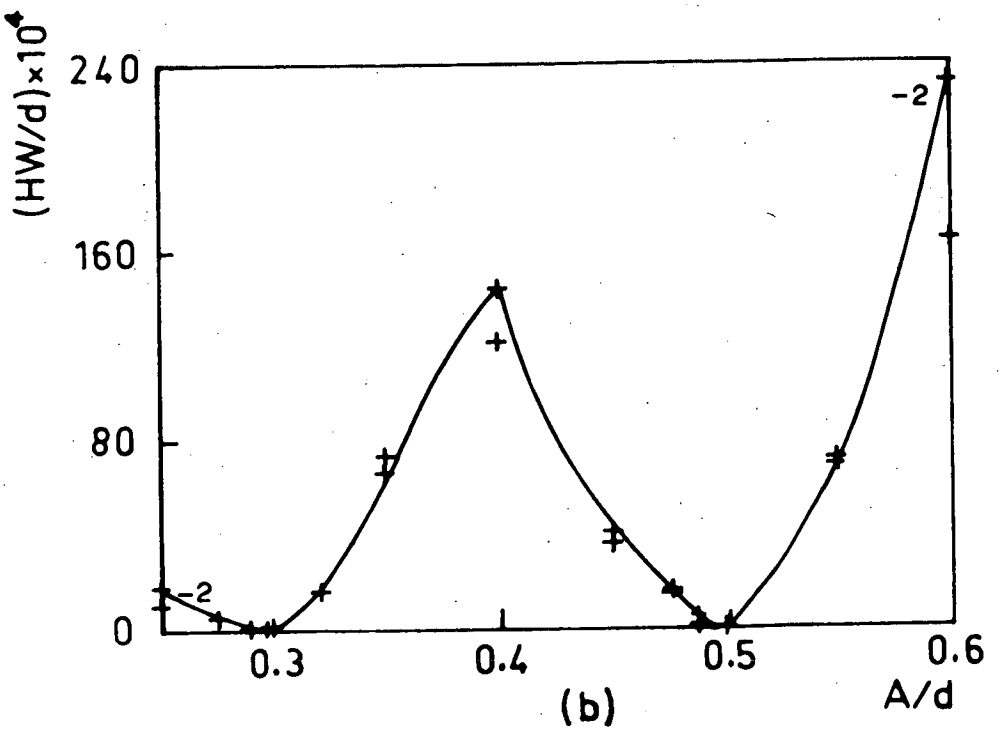
It can be seen from Figure (4.7) that the variation of ES/EP with A/d is not correlated with the variation of the strength of the P anomaly in order -1 with amplitude. This is not surprising, since the value of ES/EP is governed by the factors affecting S polarization behaviour as well as those affecting P polarization behaviour. It has been shown that in fact the strength of this P anomaly is linked to the absolute efficiencies of the grazing orders in P polarized light.

4.5. HIGH-ORDER S POLARIZATION RESONANCES

As the groove depth of the sinusoidal grating increases, the S polarization resonance peaks in the evanescent orders widen and start to move away from the Rayleigh wavelength (see Chapter 3). The intensity curves for the resonant orders -2 and +1 show two peaks for A/d = 0.25. The higher wavelength peak is the wider of the two, and is associated with the first S polarization resonance. The second peak lies close to the Rayleigh wavelength. It is the behaviour of this sharp peak which we will now consider.

The variations of three parameters characterizing the peaks in orders -2 and +1 are shown in Figure (4.8), as a function of groove depth. The maximum intensity of the peaks increases rapidly and passes through a very sharp maximum, which in turn is followed by a minimum and another strong maximum. The half-widths of both peaks become very small (of the order of 0.3 - 0.5 $\mathring{A}$) in the regions where the intensity is largest. The peaks tend to move away from the Rayleigh wavelength with increasing groove depth.


log I(MAX)
4.0
3.0
2.0
1.0
-2
-2
0.3
0.4
0.5
0.6
A/d
(a)


(HW/d)×10⁴
240
160
80
0
-2
-2
0.3
0.4
0.5
0.6
A/d
(b)

These variations are all analogous to those undergone by the first resonance peak in order +1 as the grating amplitude tends to zero. It is for this reason that we will term the intensity peaks S polarization higher-order resonance maxima. The two peaks in the curves of log I_{max} against A/d will be said to be due to the second and third S polarization resonances.

For the second S resonance, A/d = 0.296 ± 0.004. The third S resonance is strongest for A/d = 0.492 ± 0.004. Calculations have also been performed which indicate the presence of a fourth S resonance at A/d = 0.694 ± 0.004.

The extremely rapid intensity variations in orders -2 and +1 are accompanied by sharp efficiency changes in the propagating orders -1 and 0. For example, for A/d = 0.296, at the peak of the second resonance, the efficiency changes from 95% to 9% in only 0.4 Å.

The higher-order resonances give rise to extremely large values for the intensity of the evanescent orders, and thus are associated with convergence difficulties in the integral equation calculation of grating efficiency. This loss of accuracy may be attributed to truncation errors. In order to preserve an accuracy of better than 1%, the efficiency curves shown in Figure (4.9) correspond to amplitudes slightly removed from the resonant values. Also, the maximum intensity values indicated in Figure (4.8) cannot be guaranteed to be accurate to within 1%.

The S efficiency curve for order -1 shows both a Wood anomaly and the double anomalous feature due to the third resonance. This form of a sharp maximum and a sharp minimum linked by an extremely steep section of the efficiency curve is reproduced near the second and fourth resonances. For the second and third resonance anomalies the

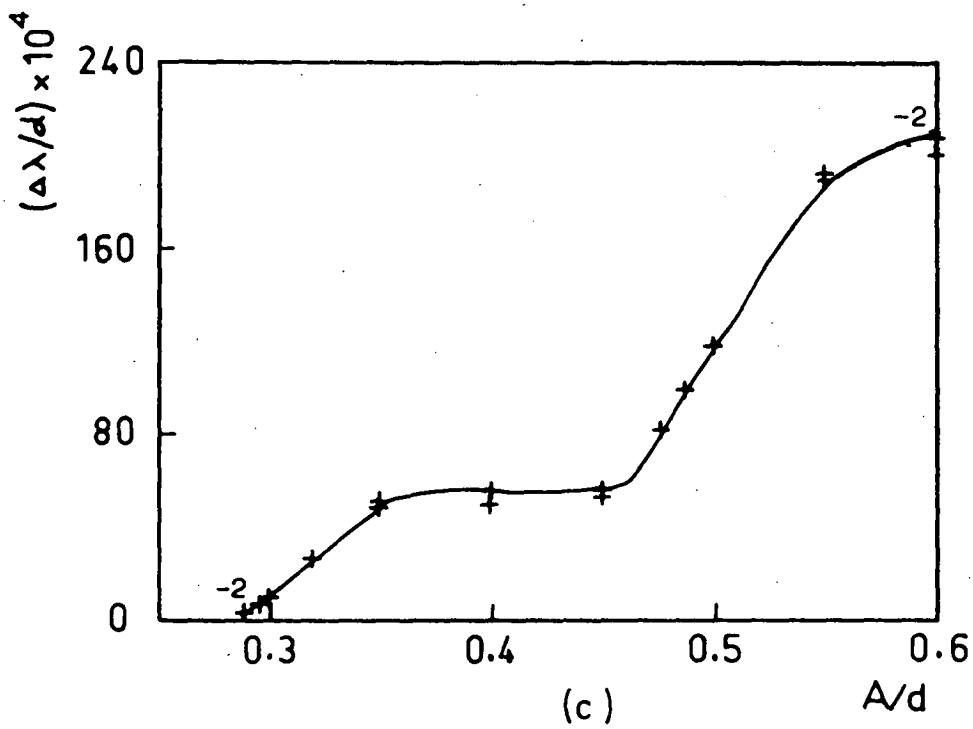


Figure (4.8). The variations of three parameters characterizing the higher-order resonance peaks in the evanescent orders -2 and +1 are shown as functions of the normalized amplitude of the sinusoidal grating profile. The curves correspond to the order -2, and the isolated points to the order +1. The ordinates are, respectively, (a) the logarithms of the maximum intensities of the resonance peaks, (b) their normalized half-widths (multiplied by 10^4), and (c) their normalized wavelength separations from the Rayleigh wavelength (multiplied by 10^4).

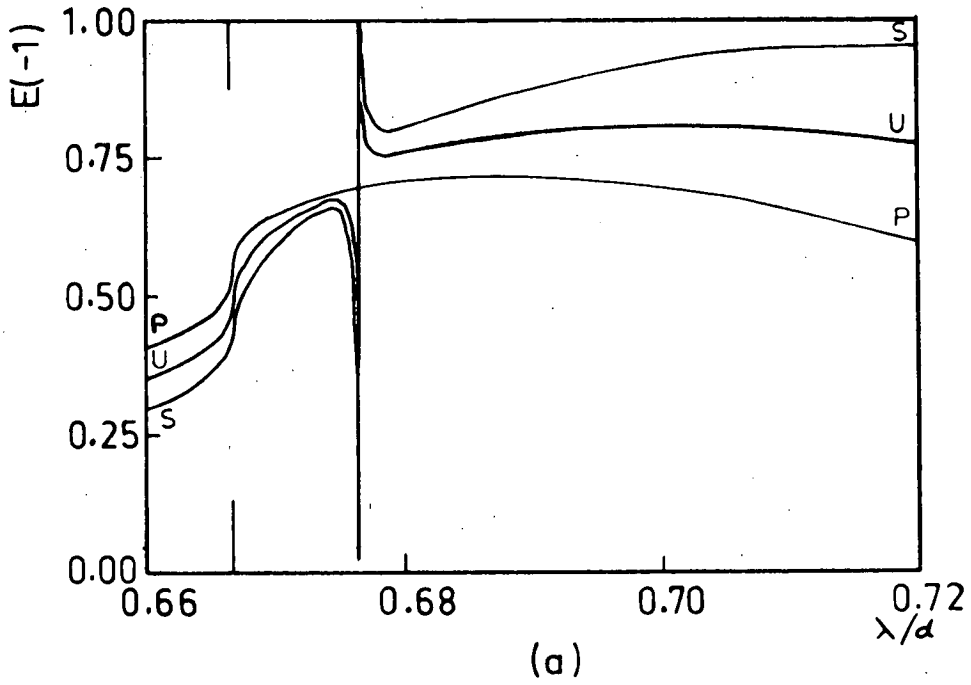
wavelength of maximum intensity for the evanescent orders lies close to the wavelength of minimum efficiency for order -1. For the fourth resonance, the wavelength of maximum intensity is associated with the efficiency maximum in order -1.

The P efficiency curve for order -1 shows only a single anomalous feature, located at the Rayleigh wavelength, and having a similar form to the S Wood anomaly. No P polarization reaction is evident in the neighbourhood of the S resonance anomaly.

The zero order efficiency curves for A/d = 0.686 are quite remarkable. Note how far the resonance anomaly has moved away from the Rayleigh wavelength. There is no perceptible S polarization Wood anomaly, but the P Wood anomaly is quite strong. This is an example of the situation observed by Palmer (4.6), but described as very unusual, of a P anomaly being formed with no related S Wood anomaly. Also unusual is the breaking of the well-substantiated association (4.6, 4.14, 4.15, 3.1) between the Rayleigh wavelength and the point of steepest gradient on the S efficiency curve.

The high-order S resonances differ fundamentally from the first resonance in one aspect of their behaviour. The first resonance peak broadens as the groove depth increases from zero and moves away from the Rayleigh wavelength. This peak is only sharp for shallow profiles. The second resonance peak also moves away from the Rayleigh wavelength with increasing groove depth, but oscillates in sharpness and strength as it does so. This peak is capable of undergoing multiple resonances, whereas the first resonance peak is not.

The oscillatory behaviour of the S polarization multiple resonances can also be interpreted in terms of the variation of the equivalent surface impedance structure (4.9). If this impedance



(a)

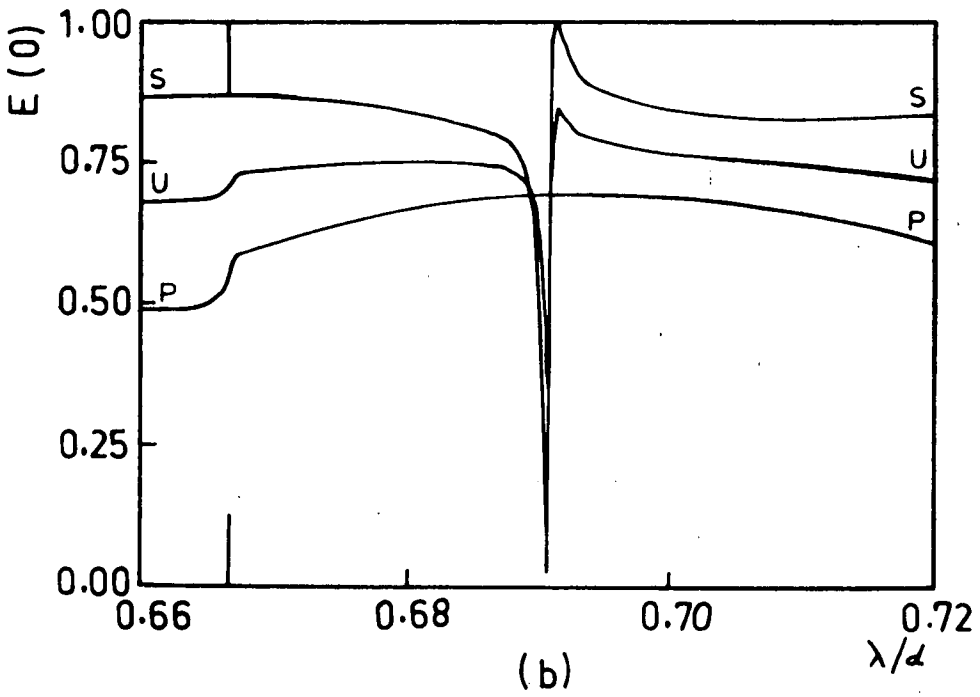


Figure (4.9). Efficiency curves for P, S and U polarized light are shown for sinusoidal gratings used in a Littrow mount in the order -1. The normalized grating amplitudes are, respectively, (a) 0.4875 and (b) 0.686.

structure is to be inductive, then the groove depth D of the sinusoidal grating must satisfy the inequality

$$\frac{n\lambda_g}{2} < D < (2n+1)\ \frac{\lambda_g}{4}\ , \quad n = 0, 1, 2, 3, \ldots .$$

Thus we would expect the strength of S anomalies to fluctuate as the groove depth moves in and out of the regions in which the above inequality holds. If we use the value $\lambda_g/d = 0.790$, then the ends of the inductive regions corresponding to n = 1,2,3 occur at the respective normalized groove depths 0.593, 0.988 and 1.383. The normalized depths at which the higher order S resonances are strongest are 0.592 ± 0.008, 0.984 ± 0.008, and 1.388 ± 0.008. These resonances then correspond to the ends of the inductive regions.

The normalized groove depths at which higher order resonances occur are thus given by the formula

$$(D/d) = [(2n+1)/4]\ (\lambda_g/d)\ , \quad n = 1, 2, 3, \ldots$$

where $\lambda_g/d = 0.790$.

4.6. CONCLUSIONS

The integral equation method of calculation of grating efficiencies has been shown to give very good agreement (in some cases) with experimental observations of P anomalies made at millimetre wavelengths. The agreement is less satisfactory for S anomalies. However, for S polarization the need for a higher angular resolution than that of the millimetre-wave spectrometer used by Palmer and Le Brun (4.8) is evident. High angular resolution is also needed for S efficiency calculations. The series-expansion formulism of

Jovicevic and Sesnic (4.7) has been shown to be less satisfactory in this respect than the integral equation method. The latter method has also been shown to have a greater angular range of validity than the modified Kirchhoff calculation recently developed by Palmer and Le Brun (4.8).

It has been demonstrated that for triangular profile gratings significant P anomalies can be formed in off-blaze orders for wavelengths up to four times the groove depth. However, the wavelength must be of the same order as the depth to give anomalies in the blaze order.

Sinusoidal profile gratings have been shown to give P anomalies which oscillate in strength as the groove depth is increased. The oscillations have been linked with the formation of first and second P resonance peaks in evanescent orders.

The existence of higher order resonance anomalies in S polarized light has been demonstrated. The second, third and fourth resonances have been shown to occur at groove depths given by a simple formula, based on an equation of Hessel and Oliner (4.9). This formula will enable the prediction of the depths for which higher resonances than the fourth occur.

The high-order resonance anomalies have been shown to give extremely rapid and large changes of efficiency with wavelength. These properties may perhaps be able to be exploited to practical advantage in spectroscopy. For this reason, experimental investigation of these anomalies is warranted. Such an experimental investigation would provide an extremely critical test of the accuracy of the theory.

REFERENCES

(4.1) McPhedran, R.C., and Waterworth, M.D., Optica Acta, in press.

(4.2) Wood, R.W., 1902, Lond.Edin.Dubl.Phil.Mag., 4, 396.

(4.3) Rayleigh, Lord, 1907, Proc.R.Soc.A, 79, 399.

(4.4) Fano, U., 1938, Annln.Phys., 32, 393.

(4.5) Artmann, K., 1942, Z.Phys., 119, 529.

(4.6) Palmer, C.H., 1952, J.opt.Soc.Am., 42, 269.

(4.7) Jovicevic, S. and Sesnic, S., 1972, J.opt.Soc.Amer., 62, 865.

(4.8) Palmer, C.H. and Le Brun, H.W., 1972, Appl.Optics, 11, 907.

(4.9) Hessel, A. and Oliner, A.A., 1965, Appl.Optics, 4, 1275.

(4.10) Pavageau, J. and Bousquet, J., 1970, Optica Acta, 17, 469.

(4.11) Petit, R., 1967, Optica Acta, 14, 301.

(4.12) Peters, C.W., Deibel, P.V., Pursley, W.K. and Zipf, T.F., 1953, Report No.2 (Engineering Research Institute, University of Michigan, Ann Arbor).

(4.13) Palmer, C.H., 1963, from "A Symposium on Electromagnetic Warfare", the Johns Hopkins University, Baltimore, Maryland.

(4.14) Stewart, J.E. and Gallaway, W.S., 1962, Appl.Optics, 1, 421.

(4.15) Ingersoll, L.R., 1921, Phys.Rev., 17, 493.

CHAPTER 5

BLAZE OPTIMIZATION FOR TRIANGULAR PROFILE GRATINGS

This chapter is based on a paper published in Optica Acta by the author and Dr. M.D. Waterworth (5.1).

5.1. INTRODUCTION

The question of the best choice of the profile of a diffraction grating from the point of view of achieving as high a diffraction efficiency as possible over its wavelength region of use is obviously one of great practical interest. Despite this, no rigorous theoretical investigation has yet been published of the connection between the profile of a grating and its spectral performance, for a wide range of grating profiles and periods. Such an investigation is reported here.

The first recommendation concerning the proper choice of grating profile was made by Trowbridge and Wood (5.2). They speculated that, on theoretical grounds, a groove with a 90^{o} apex angle should give the best performance. However, they were using a theory based on Fresnel diffraction of the incident wave, which can only be valid if the width of the groove considerably exceeds the wavelength of the radiation. This condition does not generally hold for spectrographic diffraction gratings.

G.W. Stroke (5.3) advocated the use of larger groove apex angles than 90^{o}, for instance angles of 110^{o} and 120^{o}. He considered that the important step in achieving good performance was to avoid polarization of the diffracted light, which he showed theoretically to be impossible for a rectangular profile at its blaze position. He reported that he was able to decrease the polarization of the diffracted

light and increase the efficiency of the grating at its blaze wavelength by increasing the groove apex angle to 110^{o} or 120^{o}.

In this chapter,we will consider theoretically this question of the optimal choice of profile. The criterion we will adopt is that the grating efficiency should exceed a prescribed value over as wide an unbroken wavelength region as possible. It will be assumed that polarization effects due to the grating are not detrimental to its experimental use (i.e., either the effects are unimportant to the user, or they can be allowed for). Hence, no importance will be placed on reducing the polarization of diffracted light. It will also be assumed that effects due to the variation of grating efficiency with wavelength can be taken into account, so that it is not necessary to achieve constancy of efficiency within some specified range over the grating's region of use. Thus, it is considered that the grating is to be used in a double-beam instrument, or that its reflectance as a function of wavelength will be measured. Stewart and Gallaway (5.10) and Breckinridge (5.11) have pointed out that serious errors can arise when gratings are used for intensity measurements without one or the other of these precautions being taken. Only the first order blaze of the grating will be considered, and the Littrow configuration of incident and diffracted beams will be used throughout.

Within these limitations, it will be shown that in general the best grating performance over a wide wavelength range can be secured by using small values of the groove apex angle (i.e., values not in excess of 90^{o}) rather than the larger angles recommended by Stroke.

We will make reference briefly to the theoretical method used for the calculation of grating efficiency, and give a definition of the quantity 'blaze width', before considering in detail the effect of

grating profile on spectral performance for six values of the ratio of blaze wavelength to period. We will then discuss the dependence of the wavelength shift due to polarization on grating period and profile, and will exemplify a new theoretical phenomenon. It will be suggested that further practical investigation of blaze optimization is warranted.

5.2. 'BLAZE WIDTH'

The theoretical method used for the calculation of grating efficiencies has already been described in Chapter 2 and Appendix I. We will stress only that it is assumed that the surface of the grating is infinitely conducting, and that the incident beam of radiation lies in the plane of the groove-section (i.e., that the grating is used in a two-dimensional mounting).

We will use the same notations for the polarization of radiation here as in Chapters 3 and 4. If the electric vector of the radiation is perpendicular to the rulings of the grating, the radiation is said to be S polarized. The orthogonal polarization is denoted by the letter P. Unpolarized radiation is said to be U polarized. The efficiency in unpolarized light is the average of the separate efficiencies in S and P polarized light.

Consider now the S polarization efficiency curve shown in figure (5.1). Indicated in the figure are blaze widths corresponding to minimum efficiency levels of 50%, 60%, 70%, 80% and 90%. The blaze width associated with a prescribed efficiency value E is defined to be the widest unbroken wavelength interval in which the grating efficiency exceeds E. This interval should enclose (if possible) the prescribed blaze wavelength.

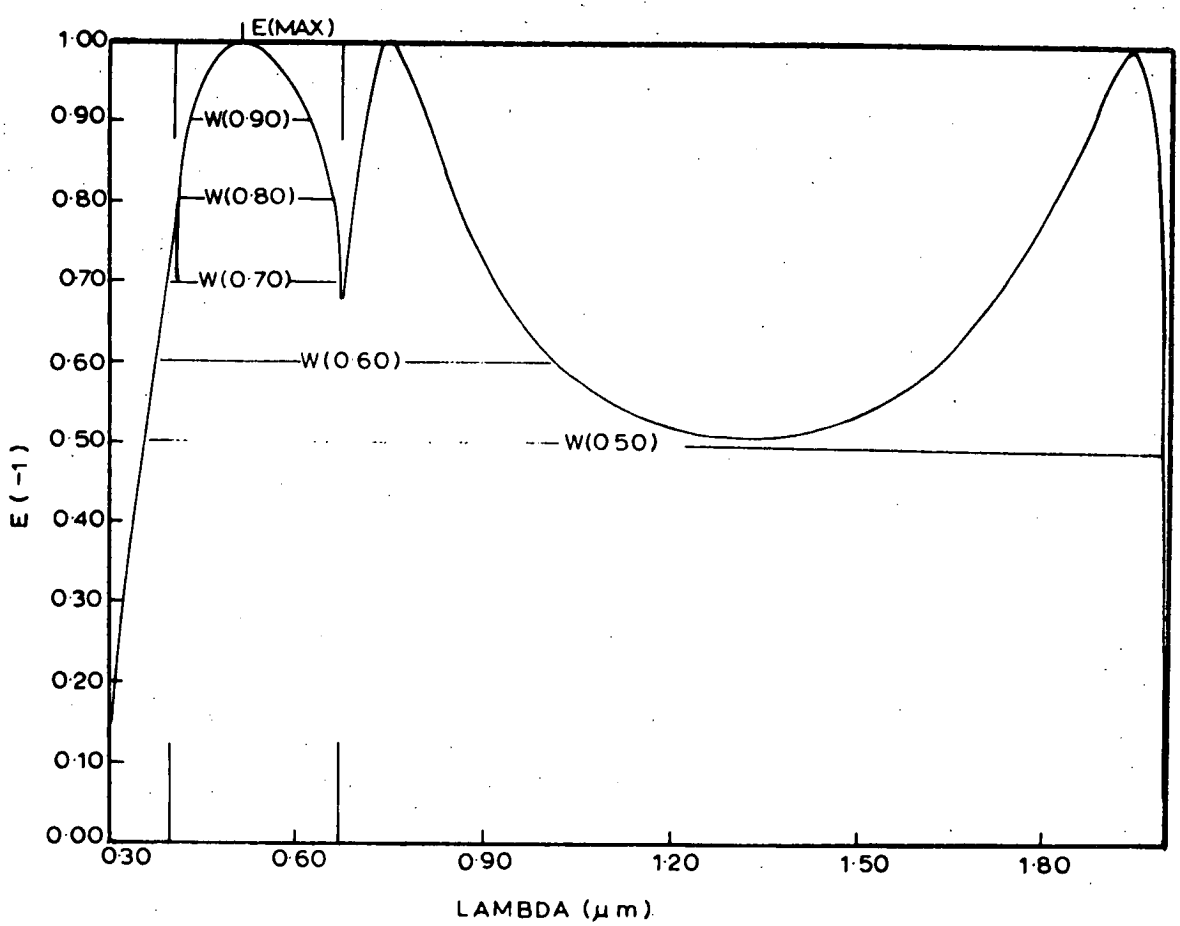


Figure (5.1). The efficiency in S polarized light for order -1 is plotted as a function of wavelength (in microns). The grating profile has a period of 1.00 μm, a blaze angle of 14.5^{0} and an apex angle of 85^{0}. Indicated in the figure are the blaze widths corresponding to efficiency cut-offs of 50%, 60%, 70%, 80% and 90%, as well as the maximum efficiency value for S polarization.

The efficiency value E associated with the blaze width is regarded as the minimum tolerable value.

With the above definition, the 60% blaze width W(0.60) does not include a contribution from the secondary blaze peak near a wavelength value of two. An alternative definition of this quantity might include as well the interval in which the efficiency exceeds 60% in this region. However, in the practical situation, some alternative dispersive device (e.g. another grating) would need to be used in the intermediate region in which the efficiency falls below the tolerable level. This device might well obviate the need for the use of the secondary blaze peak of the grating, rendering the blaze width value according to the second definition meaningless. It is for this reason that we will use the first definition of blaze width in this paper, and require the wavelength interval in which the efficiency cut-off is exceeded to be unbroken.

This definition of blaze width has the most important repercussions for S polarization. The efficiency curves for P polarization have only a single peak and a smooth shape, so that the two definitions of blaze width would be equivalent for this polarization. Also, the secondary peak is much weaker for U polarization than for S polarization (never rising much above 50%), as are the effects of Wood anomalies. (The secondary peak for U polarization, in fact, never makes a contribution to blaze width. This means that although the S polarization blaze width regions may include wavelengths for which the angle of incidence is close to 90^{o}, the corresponding U polarization regions are always associated with more moderate values of the angle of incidence.)

The S polarization efficiency curve of figure (5.1) has three strong peaks. One (at a wavelength value of 0.5) corresponds to the

main blaze of the grating. The quantity E(MAX) denotes the maximum efficiency achieved by this main blaze. The peak near a wavelength value of two corresponds to the diffracted beam lying in the position of specular reflection from the smaller, steeper groove facet, as explained by Maystre and Petit (5.4). The third peak corresponds to the resonance of the evanescent -2 and +1 orders (see Chapter 3).

In the efficiency curves shown, Rayleigh wavelengths (which give the theoretical positions of Wood anomalies) are indicated by the longer lines perpendicular to the wavelength axis.

Figure (5.2) gives efficiency curves for the three polarizations, for a grating of the same period and blaze angle as that of figure (5.1), but now with a larger groove apex angle (115^{o} instead of 85^{o}). Note the comparatively smooth form of the P polarization curve, already referred to. What was a slight dip in the S polarization efficiency curve near the Wood anomaly at 0.40000 μm has now greatly strengthened. The minimum between main and subsidiary S polarization peaks has fallen well below 50%, eliminating contributions to blaze widths from the latter. The S polarization efficiency at the blaze wavelength has decreased. It is changes such as these which shape the blaze width curves given below.

The terminology used in this paper refers to the visible wavelength region. The periods of the gratings considered range from 1/3 μm to 2.00 μm. Thus they vary from the highest line densities which holographic techniques are currently providing for the visible region (5.5, 5.6) right through the range provided by conventional ruling techniques for the same region. The blaze angles of the gratings are all chosen to give a blaze wavelength near 0.500 μm, i.e., near the centre of the visible region.

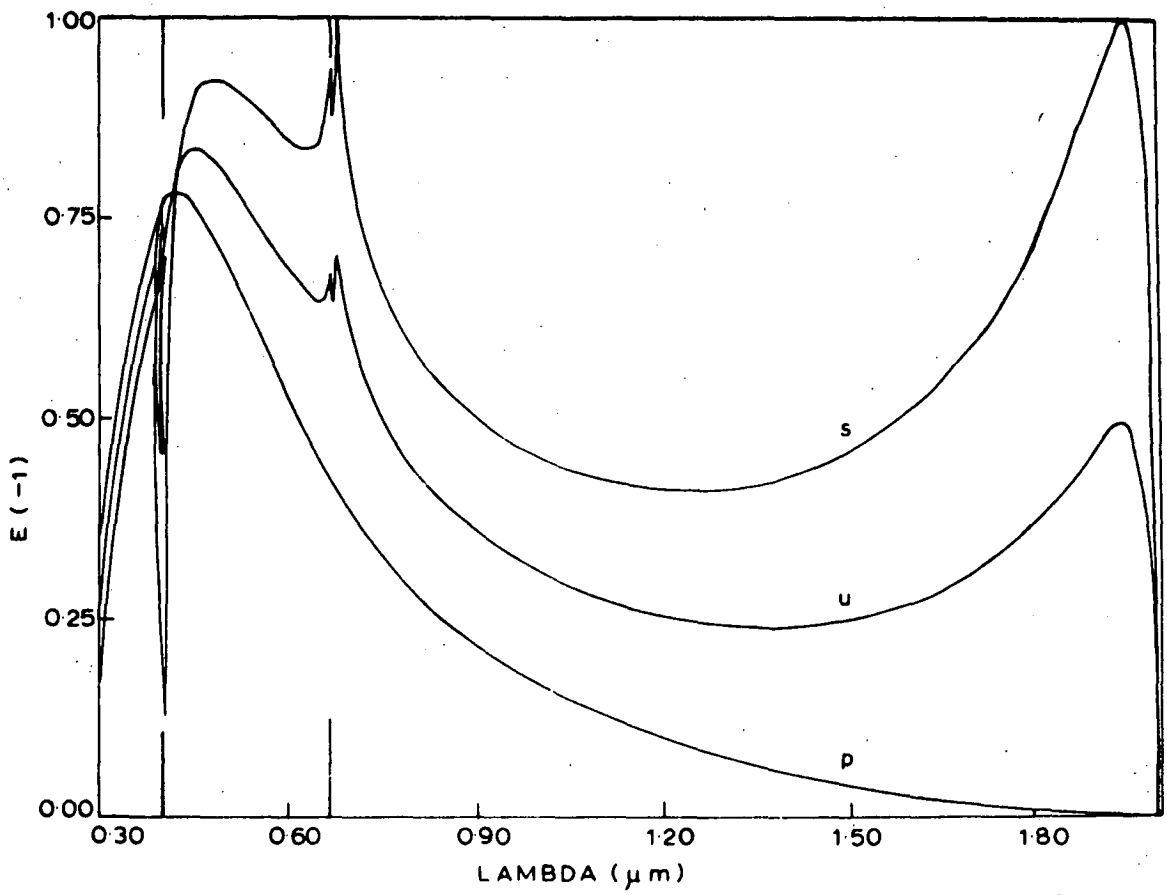


Figure (5.2). Efficiency curves are shown for P, S and U polarizations, for a grating with a period of 1.00 μm, a blaze angle of 14.5^0 and an apex angle of 115^0. Note the differences in the forms of the two S polarization Wood anomalies of figures 1 and 2.

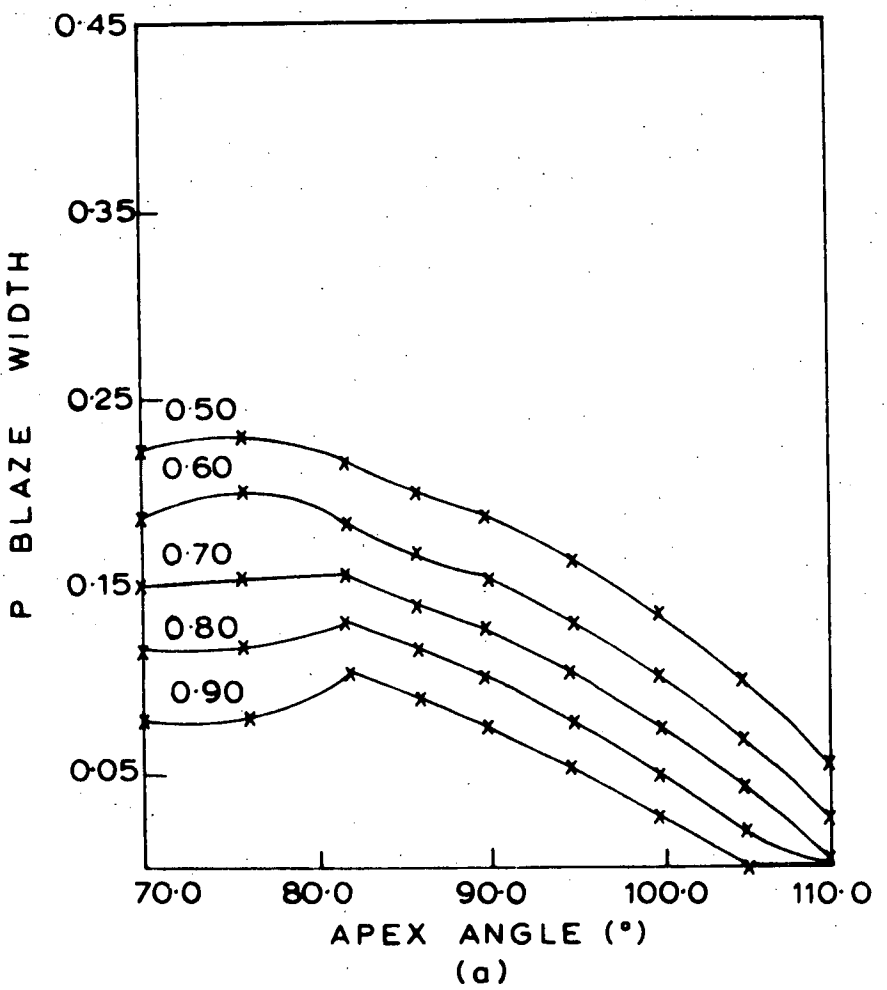
Despite this concentration on the visible region, the results are equally relevant to all wavelength regions in which echelette gratings are used. This is true since grating efficiency depends rather on the ratio of wavelength to period than on either parameter considered separately. Thus, if it is desired to rescale the wavelength region by any factor, it is only necessary to multiply the grating periods used here by the same factor.

5.3. BLAZE WIDTH AND MAXIMUM EFFICIENCY CURVES

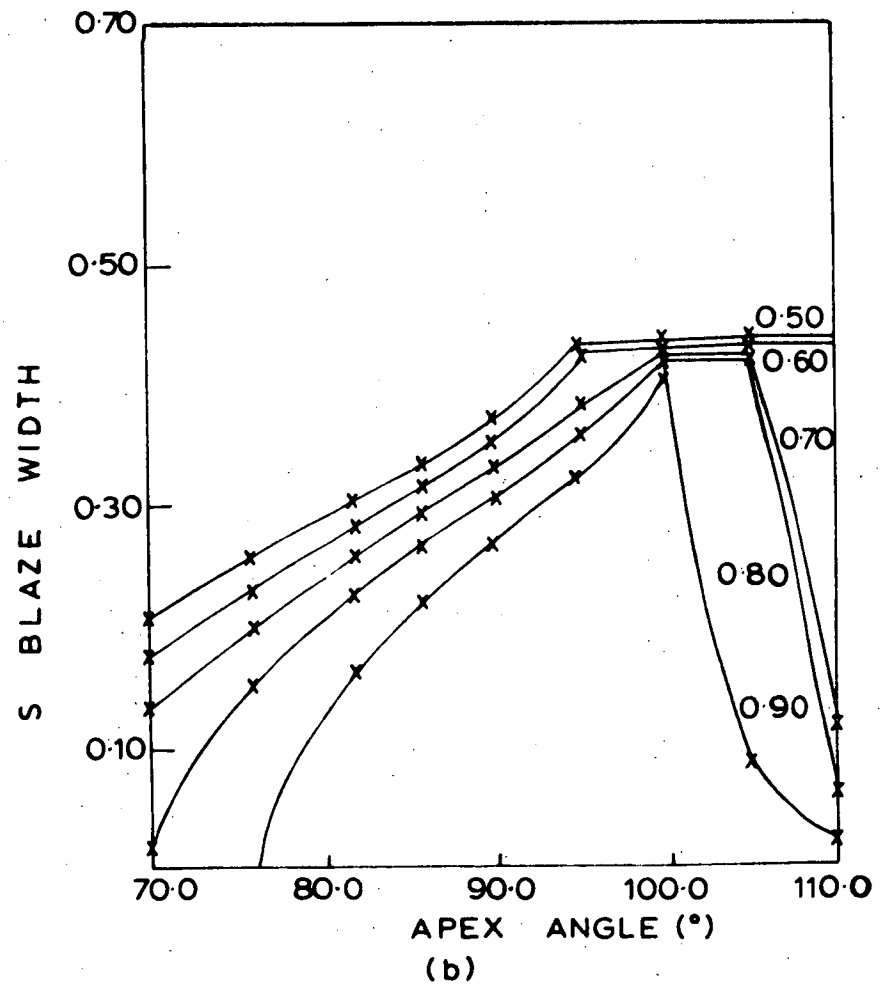
For each of the six periods discussed below, efficiency curves were calculated for a range of apex angles. On the average, each efficiency curve was characterised by values at about 110 wavelengths, with all calculations being accurate to 1% or better. From each curve, values for the efficiency maximum E(MAX) in the blaze region, and the blaze widths corresponding to 50%, 60%, 70%, 80% and 90% minimum efficiency levels were measured, for all three polarizations. The variations of these quantities with groove apex angles were displayed in graphical form.

(a) A period of 1/3 μm (Figure (5.3))

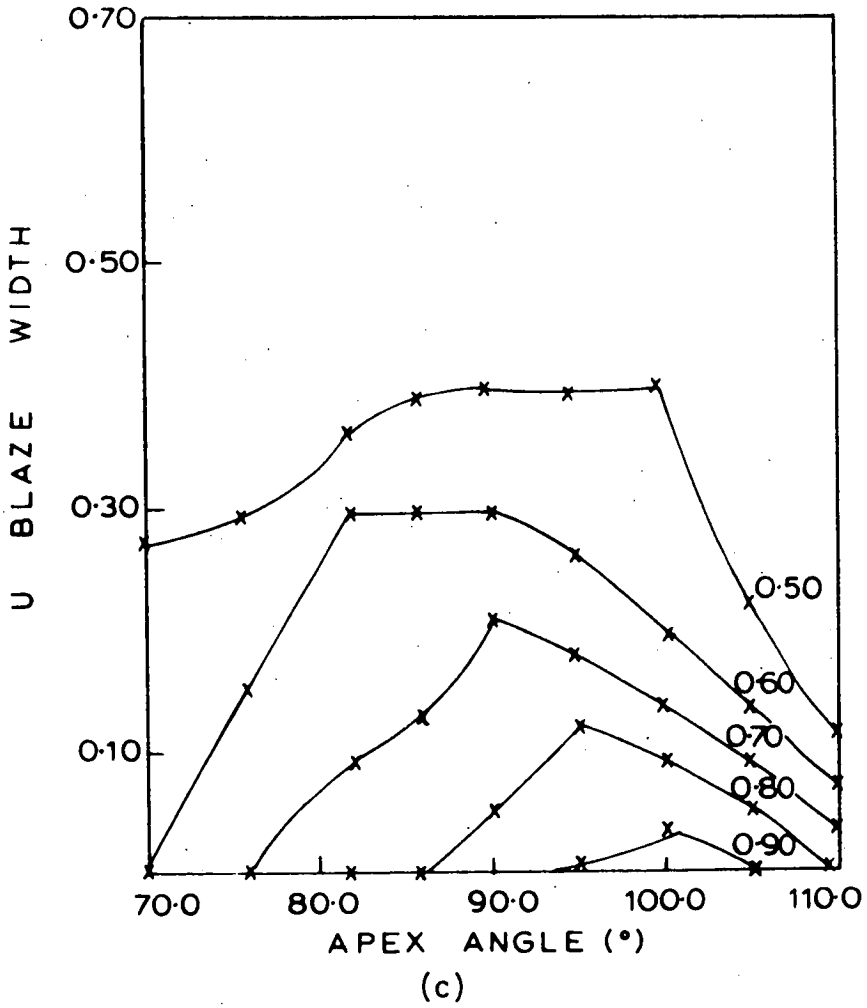
For P polarization (figure (5.3)(a)) all blaze width curves peak for apex angles lying between 76^{o} and 82^{o}. Above 82^{o}, the curves fall away quite steeply. The S polarization curves (figure (5.3)(b)) on the other hand tend to rise with increasing apex angle, until an optimum level is reached when the apex angle attains 95^{o} or 100^{o}. The sharp fall-offs in the curves corresponding to 70%, 80% and 90% levels are due to the successive isolation of the main and subsidiary blaze peaks. As the apex angle increases, the dip between these two peaks


0·45
0·35
0·25
0·15
0·05
P BLAZE WIDTH
0·50
0·60
0·70
0·80
0·90
70·0
80·0
90·0
100·0
110·0
APEX ANGLE (°)

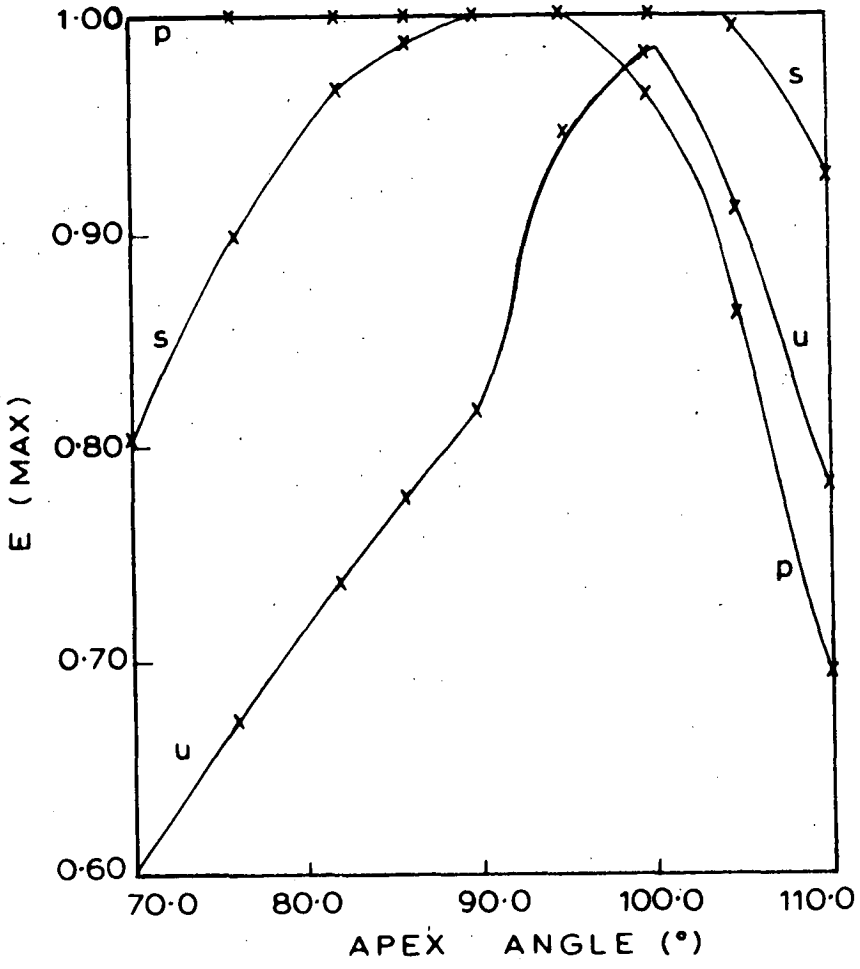
(a)


0·70
0·50
0·30
0·10
S BLAZE WIDTH
0·50
0·60
0·70
0·80
0·90
70·0
80·0
90·0
100·0
110·0
APEX ANGLE (°)

(b)


0·70
0·50
0·30
0·10
U BLAZE WIDTH
0·50
0·60
0·70
0·80
0·90
70·0
80·0
90·0
100·0
110·0
APEX ANGLE (°)

(c)



(d)

Figure (5.3). The graphs refer to gratings having periods of 1/3 μm, blaze angles of 49^0 and apex angles varying between 70^0 and 110^0. As in the corresponding figures for longer periods, efficiency cut-off levels are indicated by values placed adjacent to the relevant curve, and light polarizations are denoted by letters near the appropriate curve. The blaze width curves shown are (a) for P polarized light, (b) for S polarized light and (c) for U polarized light. Curves showing the variation of maximum efficiency values with groove apex angle for these three polarizations are given in figure (5.3)(d).

deepens, and falls below 90% for an apex angle of 105^{o}, and then below 70% for an apex angle of 110^{o}. The curves for unpolarized radiation (figure (5.3)(c)) display the resultant effects of the opposing P and S polarization trends. They all peak for apex angles lying between 90^{o} and 100^{o}, and fall away sharply for values approaching 70^{o} and 110^{o}. For efficiency cut-off values up to 70%, 90^{o} is the optimal value for the apex angle. For 80% and 90% minimum levels, the most desirable values are respectively 95^{o} and 100^{o}.

Consider now the variation of maximum efficiency values (figure (5.3)(d)). For P polarized light, E(MAX) is very close to 100% for apex angles up to 95^{o}, but then begins to drop rapidly. (For the significance of this value of 95^{o}, refer to section 5.4.) In S polarized light, the peak efficiency value is close to 100% for apex angles between 90^{o} and 105^{o}, but falls away steeply on either side of this range. In unpolarized light, the optimum choice of apex angle to give as high a peak efficiency value as possible is 100^{o}.

(b) A period of 0.5 μm (Figure 5.4))

For P polarization all blaze width curves decrease monotonically as the groove apex angle increases from 70^{o}. The curves become steeper with increasing apex angle, but the rates of decrease are smaller than for the previous period. The blaze width curves for S polarized light show less sensitivity to apex angle changes for this period than for the previous one. The only curve to show marked profile sensitivity is that corresponding to a minimum efficiency of 90%. This drops away abruptly from its peak at 100^{o}, when the blaze and subsidiary maxima become separated by a dip falling below 90%. The other curves rise gradually towards peaks at 105^{o} (80% curve) or beyond

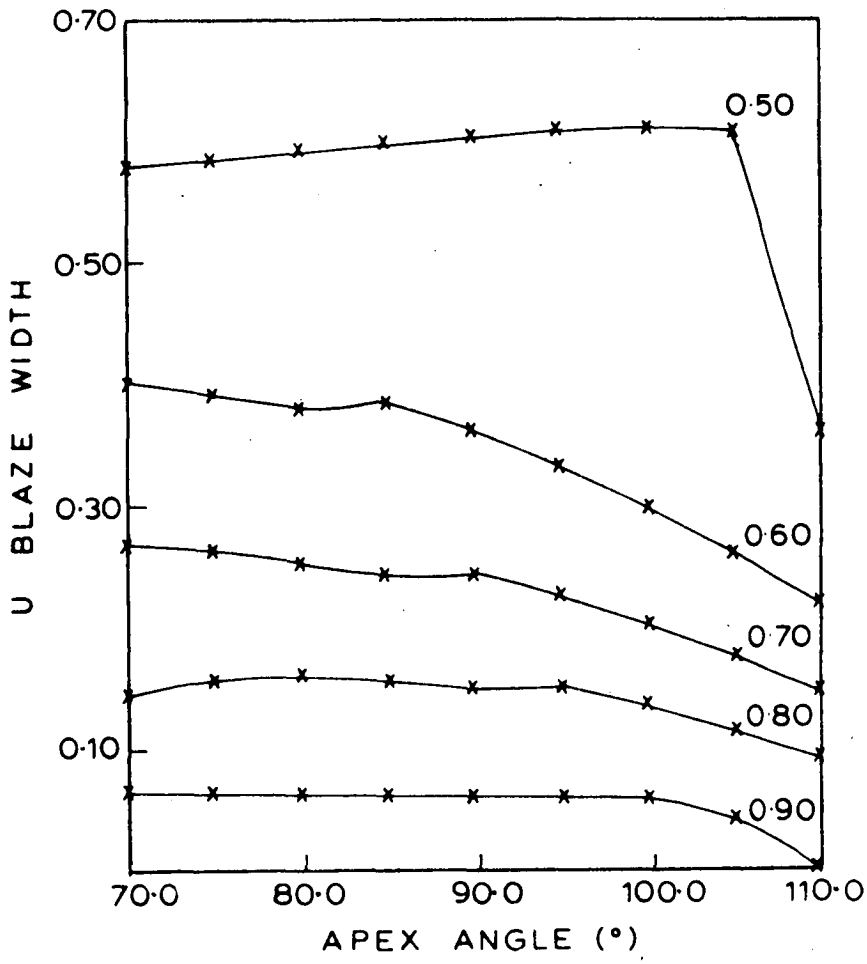


Figure (5.4). This graph refers to gratings having periods of 1/2 μm, blaze angles of 30^0 and apex angles varying between 70^0 and 110^0. The blaze width curves shown are for unpolarized light.

110^{o} (50%, 60%, 70% curves). In unpolarized light (figure (5.4)), the profile sensitivity is also less marked for this period than for the previous one. The P polarization behaviour predominates over that for S polarization for minimum efficiency levels above 50%, in the sense that the U polarization curves for 60%, 70% and 90% levels increase as the apex angle tends towards lower values, and the 80% curve peaks at a value of 80^{o}. On the other hand, the 50% blaze width curve is influenced predominantly by S polarization behaviour, since it increases gradually towards a maximum at 100^{o}, and falls off abruptly for apex angles in excess of 105^{o}.

The peak efficiency value in P polarized light is equal to 100% for apex angles smaller than 85^{o}, and falls away smoothly as the angle increases from this value. The peak efficiency value in S polarized light is always close to 100%. For unpolarized light, the peak efficiency drops away from 100% as the apex angle is increased above 75^{o}.

(c) A period of 0.75 μm (Figures (5.5) and (5.6))

For P polarization, all blaze width curves decrease monotonically as the groove apex angle increases. All blaze width values corresponding to a minimum level of 90% are zero, as are the values corresponding to 80% for apex angles in excess of 100^{o}. The S polarization curves are particularly interesting, since they show that the previous trend of blaze width to increase with apex angle (at least until around 100^{o}) has either been reversed or severely weakened. The main and subsidiary S polarization maxima become isolated with respect to a 70% minimum efficiency level for apex angles in excess of 85^{o}. For the two smaller periods, isolation did not occur until at least

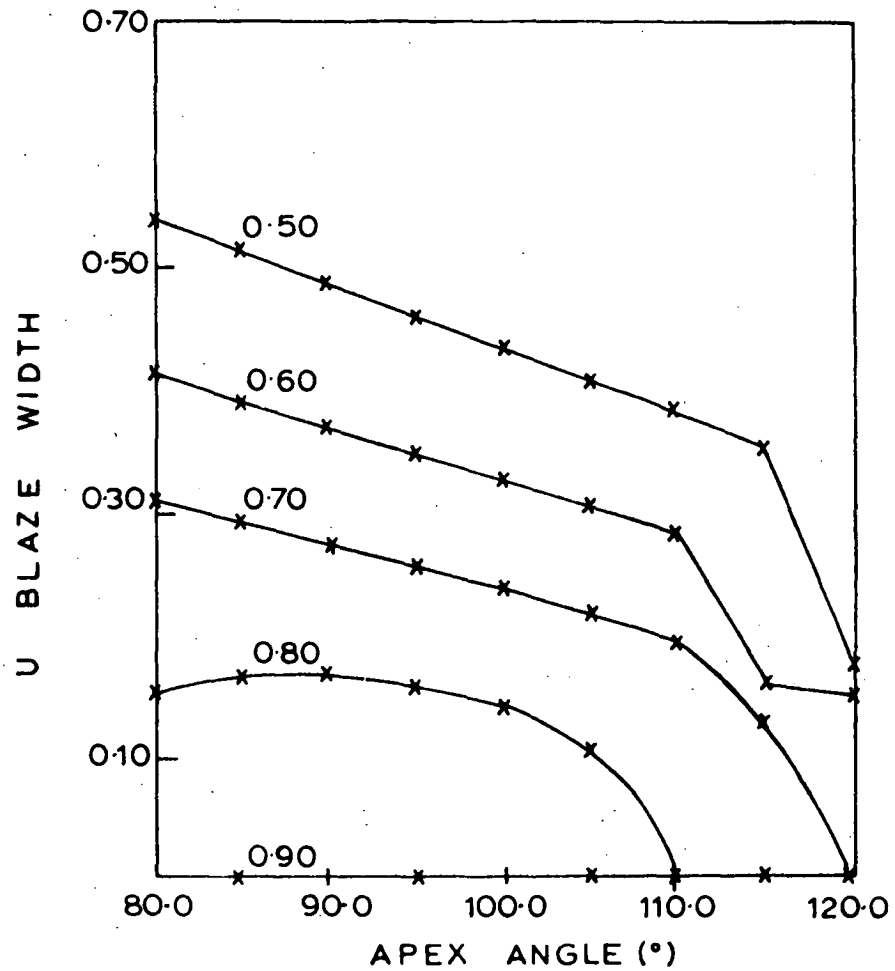


Figure (5.5). This graph refers to gratings having periods of 3/4 μm, blaze angles of 19^0 and apex angles varying between 80^0 and 120^0. The blaze width curves shown are for unpolarized light.

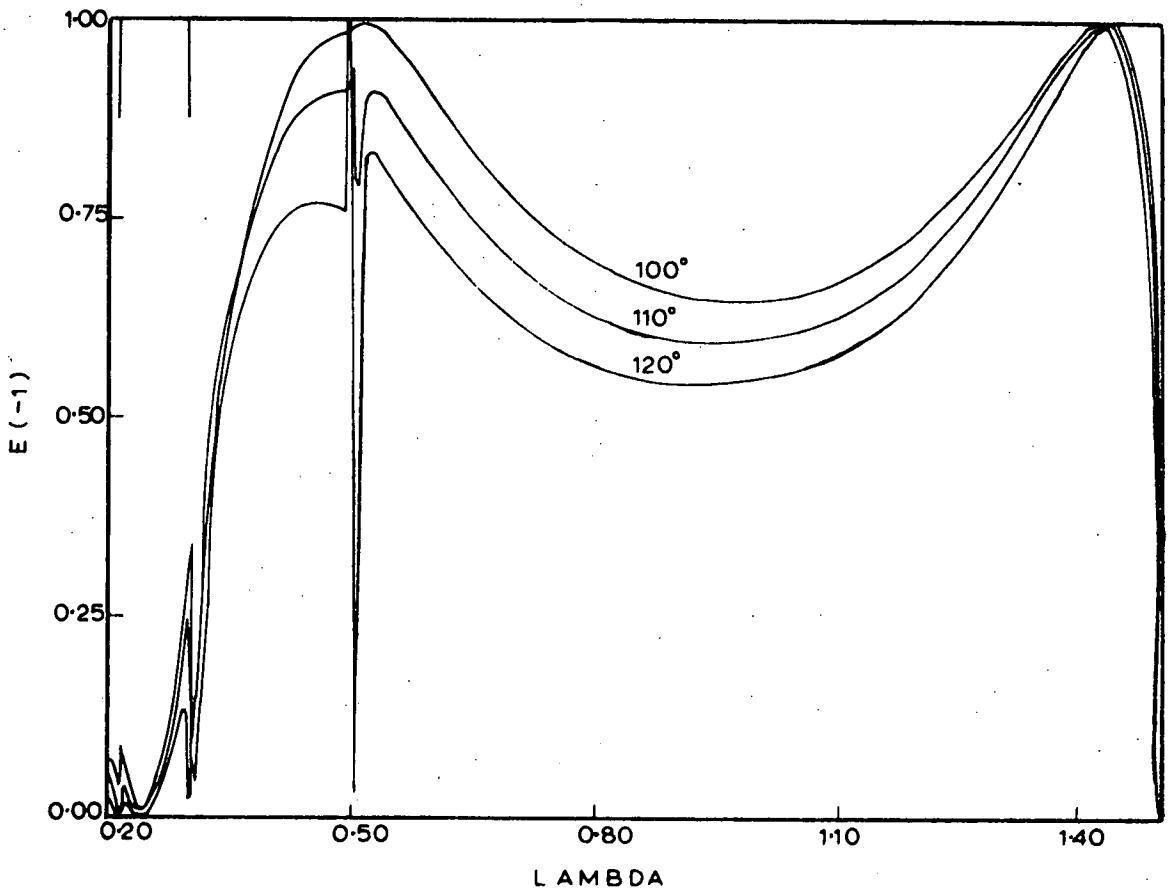


Figure (5.6). These S polarization efficiency curves are for gratings having periods of 3/4 μm, blaze angles of 19^0 and apex angles of 100^0, 110^0 and 120^0. Note how the Wood anomaly at a wavelength of 1/2 μm strengthens as the apex angle increases, from a slight bump on the 100^0 curve to a spike followed by a strong minimum on the 120^0 curve.

100^{o}. The two maxima become separated by a greater wavelength interval as the period increases, and the S polarization efficiency curve falls to ever lower values between them. W(0.60) decreases dramatically between 105^{o} and 115^{o}. Isolation of the two maxima occurs between 105^{o} and 110^{o}, but between 110^{o} and 115^{o} the Wood anomaly at 0.500 μm strengthens until it virtually cuts the main blaze maximum in half. This effect is illustrated in figure (5.6), where the S polarization efficiency curves associated with groove apex angles of 100^{o}, 110^{o} and 120^{o} are shown. (The curve for 120^{o} suggests a possible use for a grating having this profile as an S polarization notch-rejection filter. This would complement the grating arrangements devised by Maystre and Petit (5.7). for use as S polarization band-pass filters.) For the unpolarized light curves (figure (5.5)), those corresponding to 50%, 60% and 70% levels indicate optimum apex angles lying below 80^{o}, while the 80% curve peaks near 90^{o}. The effects of the S polarization Wood anomaly at 0.500 μm can also be seen in the sharp decreases between 110^{o} and 120^{o} of the first three curves mentioned above.

The efficiency maximum in P polarized light falls away smoothly and monotonically as the apex angle increases. The values are well below the corresponding values for the two previous periods. For S polarization, the efficiency maximum drops away from 100% as the apex angle increases beyond 95^{o}. The value for unpolarized light also falls with increasing apex angle. (For apex angles exceeding 100^{o}, the S and U efficiency maxima were taken to be associated with the spikes on the short-wavelength side of the anomalous minima.)

(d) A period of 1.00 μm (Figures (5.1), (5.2),(5.7))

The P polarization blaze width values again decrease

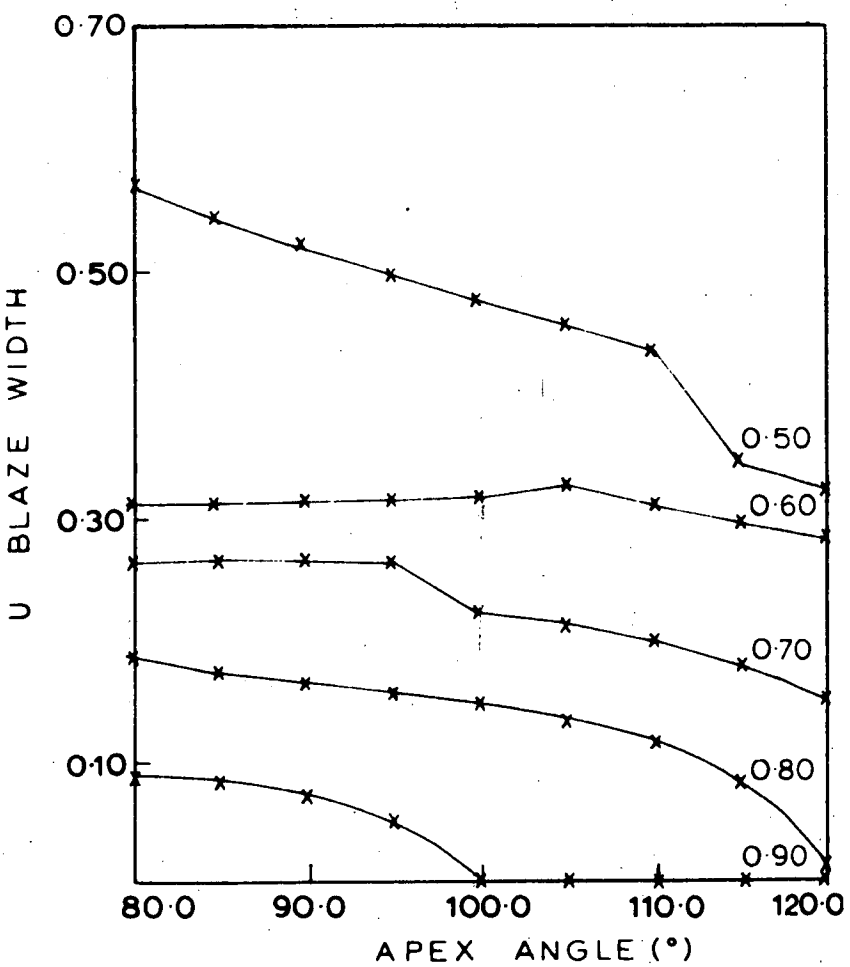


Figure (5.7). This graph refers to gratings having periods of 1 μm, blaze angles of 14.5^0 and apex angles varying between 80^0 and 120^0. The blaze width curves shown are for U polarized light.

monotonically as the apex angle increases. The behaviour of the S polarization blaze widths is complicated by the influence of two Wood anomalies, one at 0.4000 μm (a bright band on the edge of a dark band), and another at 0.66667 μm (a dark band followed by a strong resonance maximum). As the apex angle is increased, the minimum near 0.4000 μm becomes more pronounced, while the resonance maximum becomes sharper and the minimum preceding it moves to higher efficiency levels. (These changes can be observed if the S polarization efficiency curves of figures (5.1) and (5.2) are compared.) The curve corresponding to a minimum efficiency level of 50% has an abrupt drop between 85^{o} and 90^{o}, which is caused by the isolation of the main and subsidiary blazes. An increase in the curve corresponding to a 70% level between 90^{o} and 95^{o} is attributable to the movement of the minimum near 0.66667 μm above the cut-off efficiency. The minimum moves through the 80% level between 105^{o} and 110^{o}. Apart from the rises caused by the influence of Wood anomalies, the curves show a downward trend with increasing apex angle. The curves associated with 70% and 80% levels peak at 95^{o} and 110^{o} respectively, whereas the other three curves have their maxima at or below 85^{o}. Three of the curves of figure (5.7) for unpolarized light show uniform downward trends with increasing apex angle (those corresponding to levels of 50%, 80% and 90%). The 60% curve peaks near 105^{o}, while the 70% curve attains its maximum between 90^{o} and 95^{o}. The sharp decreases in the 50% and 70% curves are associated with the deepening of the minimum near 0.4000 μm. No such decrease occurs in the 60% curve since the effect of the deepening of this minimum is counter-balanced by the upward movement of the dark band associated with the second Wood anomaly.

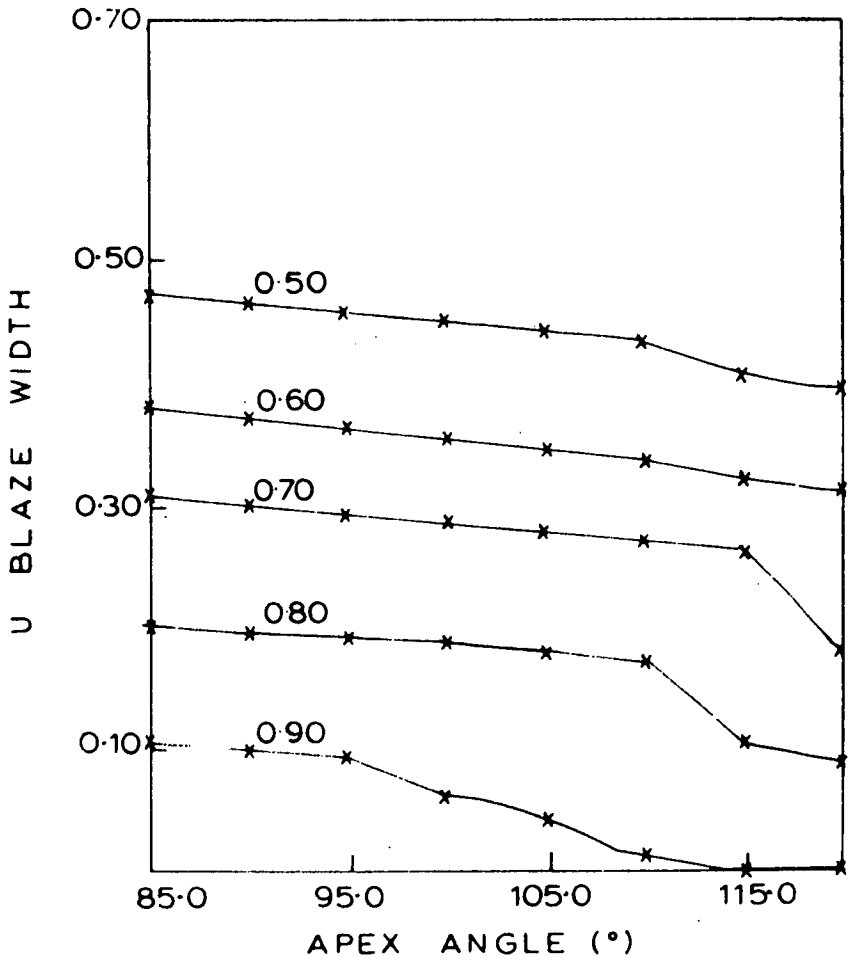


Figure (5.8). This graph refers to gratings having periods of 1.5 μm, blaze angles of 9^0 and apex angles varying between 85^0 and 120^0. The blaze width curves shown are for unpolarized light.

The efficiency maximum in P polarized light decreases monotonically with increasing apex angle, as does that in unpolarized light. The curve for S polarized light peaks at 90^{o}.

(e) A period of 1.50 μm (Figure (5.8))

The blaze width curves for P polarized light decrease monotonically as the apex angle increases. However, the curves all slope more gently down than did the corresponding curves for a period of 1.00 μm. The general trend of the S polarization curves is also to decrease with increasing apex angle. An upward jump in the 80% curve is due to the weakening of the dark band associated with the Wood anomaly at 0.6000 μm. The strengthening in the dark band of the Wood anomaly at 0.42857 μm causes downward edges which occur successively in the 90%, 80%, 70% and 60% curves. The curves in figure (5.8) for unpolarized light also show a downward trend as the apex angle increases. The downward edge in the 50% curve is associated with the dark Wood anomaly at 0.33333 μm, while those in the 70%, 80% and 90% curves are associated with the anomaly at 0.42857 μm referred to above.

The maximum efficiency curves for P and U polarized light have similar forms, both decreasing smoothly with increasing apex angle. The S polarization curve peaks for an apex angle of 90^{o}.

(f) A period of 2.00 μm (Figure (5.9))

The P polarization blaze widths decrease smoothly with increasing apex angle. The trend of decreasing profile sensitivity with period is continued. Of the S polarization curves, only that for a minimum efficiency level of 60% peaks at an apex angle in excess of

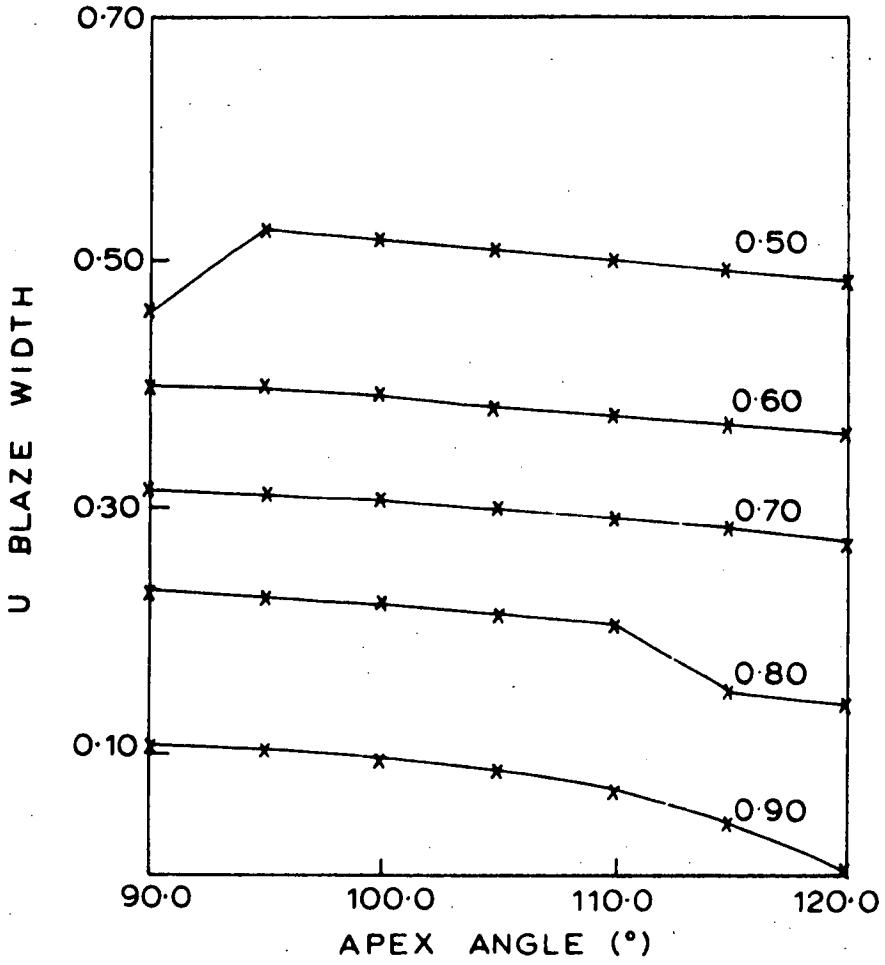


Figure (5.9). This graph refers to gratings having periods of 2 μm, blaze angles of 7.2^{0} and apex angles varying between 90^{0} and 120^{0}. The blaze width curves shown are for U polarized light.

90^{o}. The 60% blaze width is effectively constant for apex angles up to 105^{o}. The increase there is due to the weakening of the dark band associated with the Wood anomaly at 0.80000 μm. Successive downward edges in the 80%, 70% and 60% curves are associated with the strengthening of the dark band of the Wood anomaly at 0.44444 μm. Of the curves for unpolarized light in figure (5.9), all but that corresponding to a 50% efficiency cut-off optimize at or below 90^{o}. The upward step in this curve is again linked with the weakening dark band near 0.8 μm. The downward jump in the 80% curve is due to the strengthening of the dark band near 0.44444 μm.

The curves showing the variation of the maximum efficiency with apex angle for the three polarizations are all similar, decreasing smoothly as the groove profile becomes more shallow.

5.4. WAVELENGTH SHIFTS WITH POLARIZATION AND WITH APEX ANGLE

Consider for profiles with a 90^{o} apex angle the wavelength difference ($\Delta\lambda$) between the P polarization point of peak efficiency and the S polarization blaze wavelength (λ_B). Figure (5.10) shows the variation of $\Delta\lambda/\lambda_B$ with the ratio of blaze wavelength to grating period (λ_B/d). This theoretical graph should be compared with the experimental findings of Yakovlev and Gerasimov (5.8), shown in Figure 4 of their paper. Six of the ten theoretical points define a good straight line, which has the equation

$$\frac{\Delta\lambda}{\lambda_B} = \frac{\alpha\lambda_B}{d} \quad ,$$

where $\alpha = 0.344$. The experimental value of the gradient of the straight line is $\alpha = 0.34$. Despite this good agreement, it will be

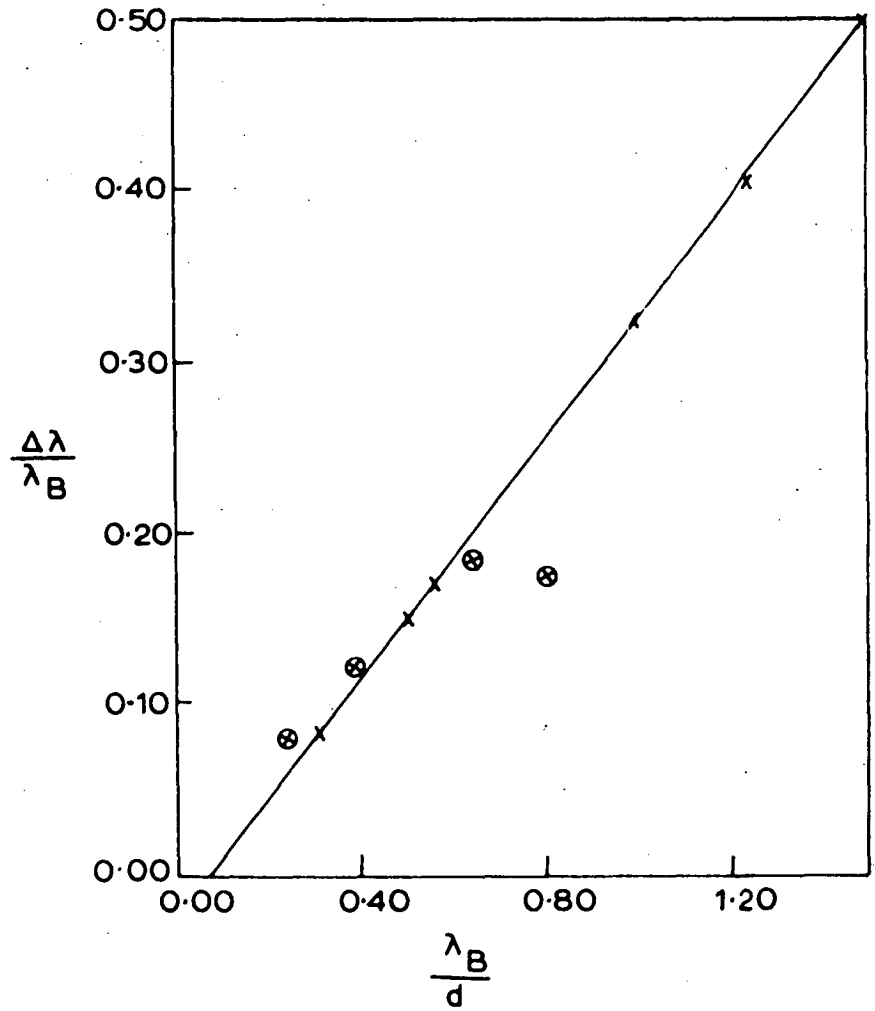


Figure (5.10). The ratio $\Delta\lambda/\lambda_B$ of the wavelength shift due to polarization and the S polarization blaze wavelength is given as a function of the ratio λ_B/d of the blaze wavelength and the period, for gratings having rectangular grooves and periods from 1/3 to 2 μm. The circled points have been affected by Wood anomalies. The gradient of the straight line shown is 0.344.

noted that the four circled points lie distinctly off the line. The two points furthest from the line have been affected by Wood anomalies which lie very close to the P polarization efficiency maxima. Wood anomalies lie close to the S polarization blaze wavelengths (λ_B) associated with the other two points. (The efficiency curves on which figure (5.10) is based are shown in Appendix II.)

Figure (5.11) illustrates in another way the influence which Wood anomalies can exert on the behaviour of efficiency maxima. It shows the variations with apex angle of the wavelengths of the P and U polarization efficiency maxima ($\lambda_M(P)$ and $\lambda_M(U)$ respectively) for gratings of period 1/3 μm. (The S polarization blazes are too wide and flat for this period to make it possible to draw a curve for this polarization.) Both curves in figure (5.11) have similar shapes. They are almost linear for low apex angles, but then begin to slope down more steeply. For apex angles larger than 95^o, $\lambda_M(P)$ is effectively constant, and just exceeds the Rayleigh wavelength (0.22222 μm) associated with the passing-off of orders -2 and +1. The P polarization efficiency maximum has been 'captured' by the Wood anomaly. The U polarization efficiency maximum is captured by the Wood anomaly at an apex angle of 100^o. (Both the angles 95^o and 100^o are also significant in relation to the variation of the respective maximum efficiency values with the apex angle, as can be seen from figure (5.3)(d).)

For apex angles in excess of 100^o, $\lambda_M(U)$ is very close to $\lambda_M(P)$. There is also an S polarization efficiency maximum lying close to the P and U maxima. This near-coincidence of the S and P polarization efficiency peaks is in keeping with the observations of G.W. Stroke (5.3), who found that the polarization of diffracted light near the wavelength of peak efficiency could be decreased by increasing

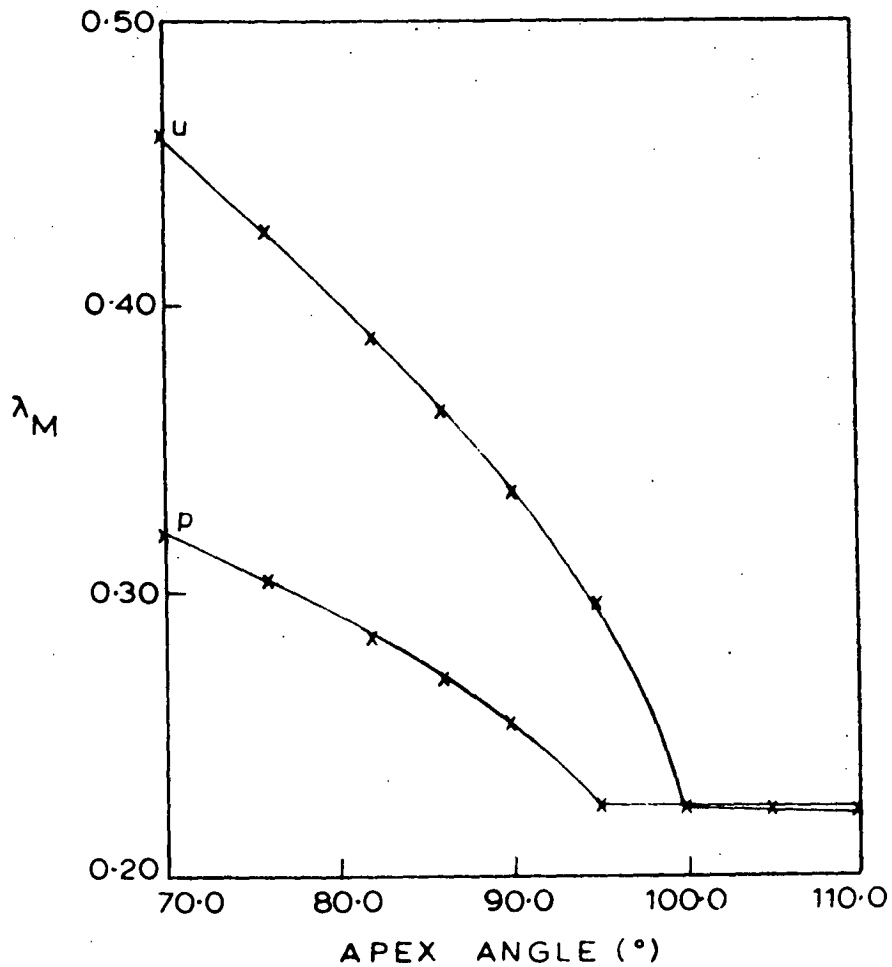


Figure (5.11). The capture of P and U polarization efficiency maxima by a Wood anomaly is shown, for a grating period of 1/3 μm. For the two polarizations, the wavelength of maximum efficiency is plotted as a function of the groove apex angle.

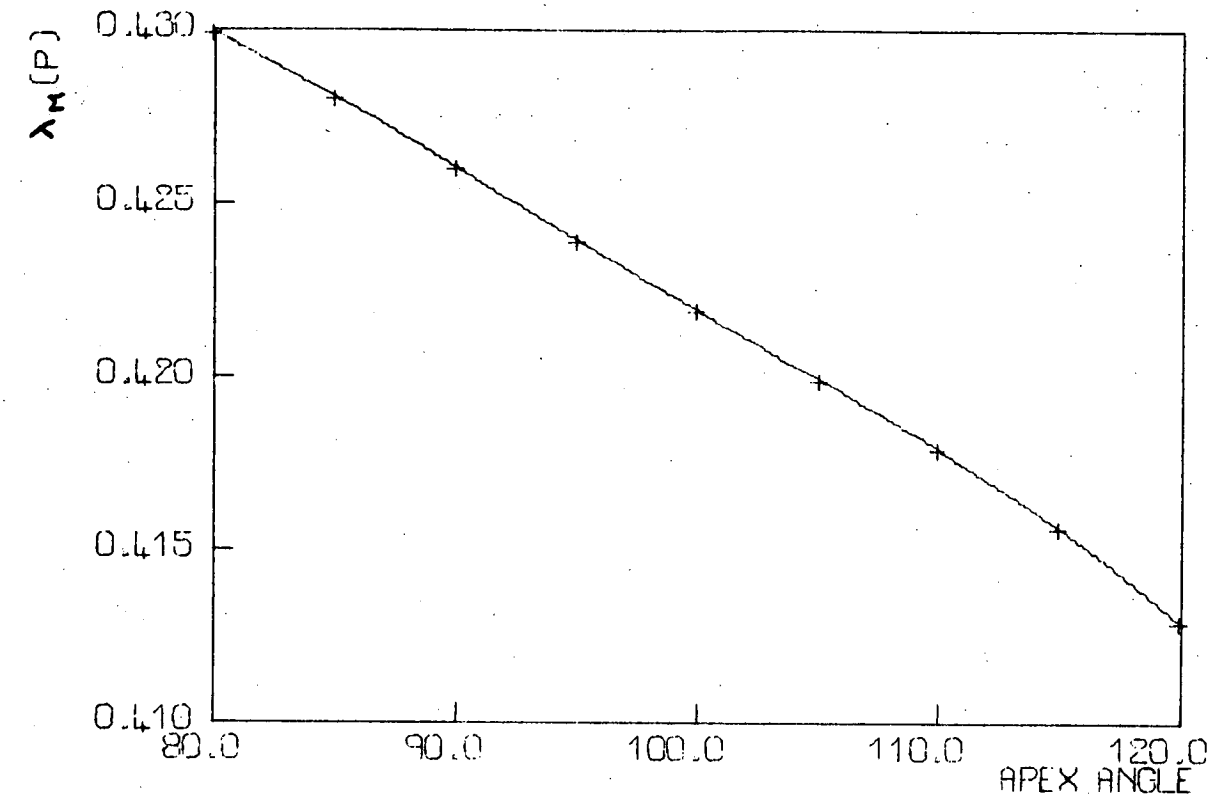
the groove apex angle from 90^{o} to 110^{o} or 120^{o}.

The capture of efficiency maxima by Wood anomalies can be understood qualitatively if we note the association between the Rayleigh wavelength and the point of steepest gradient of an efficiency curve. While the efficiency maximum lies above and close to the Rayleigh wavelength, the steepest gradient will be positive, with the efficiency curve sloping upwards towards the maximum. Were the efficiency maximum to move below the Rayleigh wavelength, a sudden reversal of the sign of the large positive gradient would be required. Such a reversal does not take place.

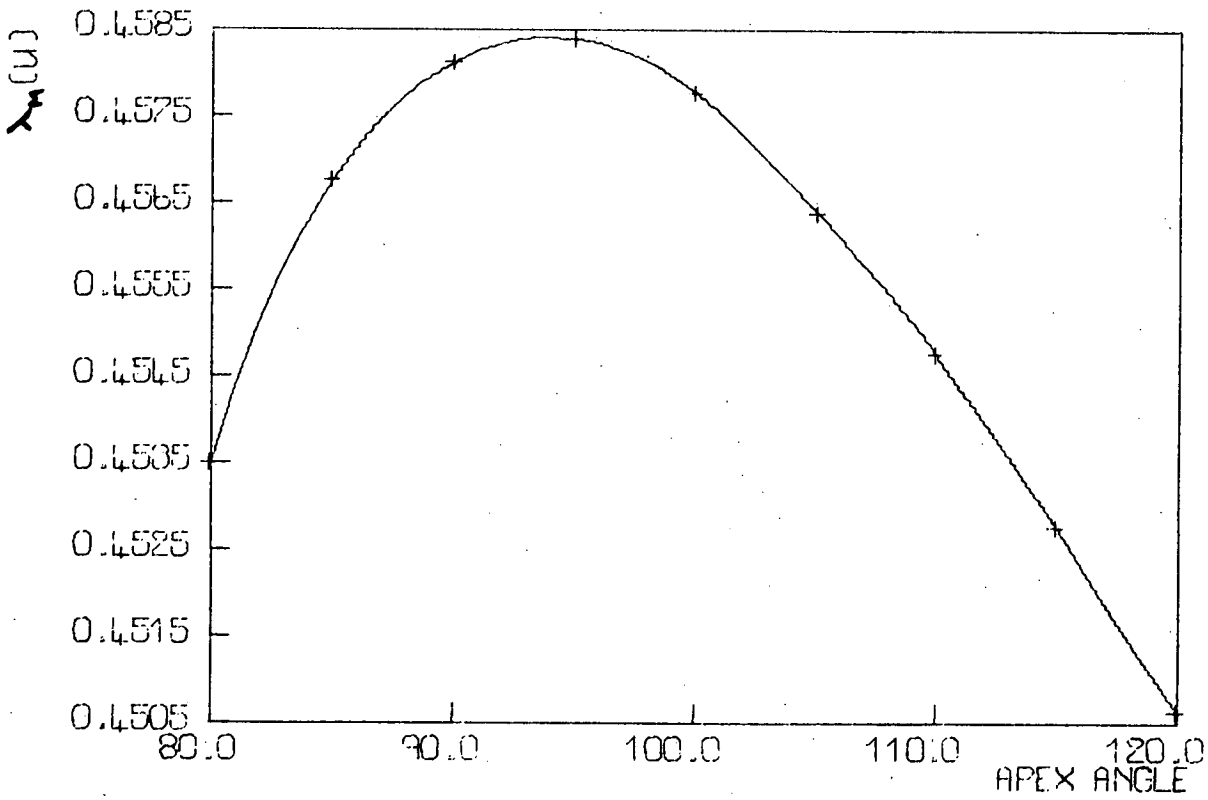
Figure (5.12) shows curves corresponding to those of figure (5.11), but now for gratings of period 1.00 μm. The P polarization points of figure (5.12)(a) define a curve which falls away linearly for apex angles up to 110^{o}, after which value the modulus of the gradient starts to increase as the efficiency maximum approaches closer to the Wood anomaly at 0.4000 μm. The gradient of the linear section is $-4\ \mathring{A}/^{o}$. This is much smaller in magnitude than is the slope ($-29\ \mathring{A}/^{o}$) between the first two points of the P polarization curve of figure (5.11). (The P polarization Wood anomaly for a period of 1.00 μm is much weaker than that for a period of 1/3 μm.)

The behaviour of the U polarization points for this period differs from that for a period of 1/3 μm. Initially, $\lambda_M(U)$ increases with apex angle. It attains its maximum value near 95^{o}, and then starts to decrease. In the section between 110^{o} and 120^{o} its decrease has become almost linear, with a gradient close to that of the linear region of the P polarization curve.

The two trends shown in the U polarization curve can be linked with the opposing influences exerted by the two S polarization



(a)



(b)

Figure (5.12). The variation of the wavelength of maximum efficiency with groove apex angle is shown, for gratings of period 1.00 μm. The top and bottom graphs refer respectively to P and U polarized light.

Wood anomalies at 0.66667 μm and 0.4000 μm. The upper anomaly consists of a dark band followed by a strong resonance maximum. The lower anomaly is a minimum, with a maximum located on its short-wavelength side. As the apex angle increases, the upper anomaly weakens, while the lower one becomes more pronounced. Thus, at 95^{o} we can say that the U polarization efficiency maximum begins to be influenced more strongly by the Wood anomaly at 0.4000 μm than by the anomaly at 0.66667 μm.

Figures (5.12)(b) and (5.11) show that the location of the U polarization point of maximum efficiency can be influenced both by S polarization and P polarization Wood anomalies.

5.5. CONCLUSIONS

(a) The P polarization blaze width curves for periods larger than 1/3 μm all decrease monotonically as the apex angle increases. For 1/3 μm, all curves peak for apex angles lying between 76^{o} and 82^{o}. Thus, if good grating performance in P polarized light is desired, the groove apex angle should be kept as small as possible, and preferably below 90^{o}.

(b) In part because of our definition of blaze width, the S polarization curves are less regular in form than the P polarization curves. The curves for a line density of 3000/mm are the only ones to show a strong trend of increasing blaze width with increasing apex angle. They reach optimum levels for apex angles around 95^{o} or 100^{o}. From a period of 0.750 μm on, the general trend is for S polarization blaze width curves to slope downwards as the apex angle increases. This is the range of

periods produced by conventional ruling techniques for use in the visible region. From the curves, it is evident that the groove apex angle should not markedly exceed 90^{o} for good S polarization performance in this range of periods.

If high S polarization efficiencies (larger than 90%) are required over a wide wavelength range, then a grating having a period of 0.500 μm and a blaze angle of 30^{o} is the ideal choice. The blaze width curves for this period show a remarkable insensitivity towards profile changes in S polarized light. (Such a grating would be well employed, for example, in providing the wavelength selection in a laser cavity (5.4, 5.9).)

(c) The grating behaviour in unpolarized light is the resultant of the separate behaviours in P and S polarized light. Hence, for periods equal to or exceeding 0.500 μm, the blaze width curves generally decrease as the apex angle increases. Groove apex angles not larger than 90^{o} are thus favoured in providing good performance for gratings having such periods. For a period of 1/3 μm, the apex angle should lie between 90^{o} and 100^{o}, as the blaze width falls away on either side of this range.

(d) Consider now the choice of profile if only the achievement of maximum efficiency at a single wavelength is of importance. If P polarized light is to be used, then regardless of the grating period, the groove apex angle should be kept as small as possible, and certainly should not exceed 90^{o}. The same conclusion is true for unpolarized light, if the line density of the grating is smaller than 3000/mm. For a grating of this line density, an apex angle of 100^{o} would be appropriate. For S polarized light,

an apex angle of 90^{o} would always be suitable, since the rectangular profile affords an efficiency of 100% at the blaze wavelength for a Littrow configuration of incident and diffracted beams.

(e) In general, the sensitivity of grating performance to groove apex angle changes decreases as the period increases. The period of 0.50 μm is exceptional in this regard, the profile sensitivity being low despite the high line density.

(f) In agreement with experimental observations, the wavelength difference between the points of maximum efficiency for P and S polarized light has been shown to vary directly as the square of the blaze wavelength and inversely as the grating period.

(g) For one period, the phenomenon of the capture of efficiency maxima by a Wood anomaly has been demonstrated.

In view of the desirability of achieving optimum spectral performance from diffraction gratings, a thorough experimental investigation of the accuracy of these theoretical predictions should be performed. Such an investigation might best be carried out in the mm wave region, where the grating profile can be determined and controlled easily.

REFERENCES

(5.1) McPhedran, R.C., and Waterworth, M.D., 1973, Optica Acta, 20, 177.

(5.2) Trowbridge, A. and Wood, R.W., 1910, Lond.Edin.Dubl.Phil.Mag., 20, 886.

(5.3) Stroke, G.W., 1963, Phys.Lett., 5, 45.

(5.4) Maystre, D. and Petit, R., 1971, Nouv.Revue Opt.Appl., 2, 115.

(5.5) Labeyrie, A. and Flamand, J., 1969, Opt.Commun., 1, 5.

(5.6) Sheridon, N.K., 1968, Appl.Phys.Lett., 12, 316.

(5.7) Maystre, D. and Petit, R., 1972, Opt.Commun., 4, 380.

(5.8) Yakovlev, E.A. and Gerasimov, F.M., 1961, Optics Spectrosc., 10, 50.

(5.9) Hard, T.M., 1970, Appl.Optics, 9, 1825.

(5.10) Stewart, J.E. and Gallaway, W.S., 1962, Appl.Optics, 1, 421.

(5.11) Breckinridge, J.B., 1971, Appl.Optics, 10, 286.

CHAPTER 6

BLAZE OPTIMIZATION FOR SINUSOIDAL PROFILE GRATINGS

This chapter is based on a paper by the author, Mr. I.J. Wilson and Dr. M.D. Waterworth (6.1), which has been accepted for publication by Optics Communications.

6.1. INTRODUCTION

The consideration of the blaze properties of diffraction gratings having a sinusoidal profile has assumed great importance since the demonstration by Brandes and Curran (6.2) that holographic methods of grating production can generate this profile. In this chapter the results of a systematic study of the blaze properties of the sinusoidal grating will be reported.

Several authors (6.3, 6.4, 6.5) have considered the problem of calculating the distribution of energy in the orders diffracted from a sinusoidal grating, using a method based on the Rayleigh expansion for the diffracted field. However, it has been demonstrated (6.6) that this method is valid only for shallow grooves.

Rigorous formulisms which enable the calculation of efficiencies for perfectly conducting gratings of arbitrary profile have been developed during the last ten years. (Summaries of these may be found in Chapter 1, and in reference (6.7).) The method of Pavageau and Bousquet (6.8), already used in Chapter 5 in a study of the blaze properties of triangular profile gratings, is applied here to sinusoidal gratings, for which some efficiency curves obtained using the conformal mapping technique have been published by M. Nevière and M. Cadilhac (6.9).

6.2. AMPLITUDE DEPENDENCE OF BLAZEWIDTH

The sinusoidal profile is specified by only two parameters, its amplitude A (or groove depth 2A) and period d. Furthermore, the efficiency of sinusoidal gratings is a function of the normalized wavelength λ/d and the normalized amplitude A/d of the profile. Thus, the results given below showing the variation of blazewidth as a function of amplitude can be applied to a grating having any desired period d' if the amplitude and wavelength are rescaled according to the formulae

$$A' = A\frac{d'}{d} \quad , \quad \lambda' = \lambda\frac{d'}{d} \quad .$$

(The blazewidth is defined, as in Chapter 5, to be the widest unbroken wavelength interval over which the efficiency exceeds a prescribed value.)

Efficiency curves have been calculated for a wavelength interval from $\lambda/d \simeq 0.20$ to 2.00 and for profile amplitudes ranging from $A/d = 0.05$ to 0.65. (The efficiency curves are shown in Appendix III.) The lower wavelength limit and the upper amplitude bound were determined by the requirement that the sum of the diffracted energy should differ from the energy of the incident wave by less than 1%. The upper wavelength corresponds to the passing off of the order -1. For all calculations the grating was assumed to be used in the Littrow mount for the order -1.

For each amplitude, the blazewidths associated with P, S and U polarized radiation and minimum efficiency levels of 60%, 70%, 80% and 90% were calculated. The variation of the most useful quantity (the blazewidth for unpolarized radiation) as a function of profile amplitude

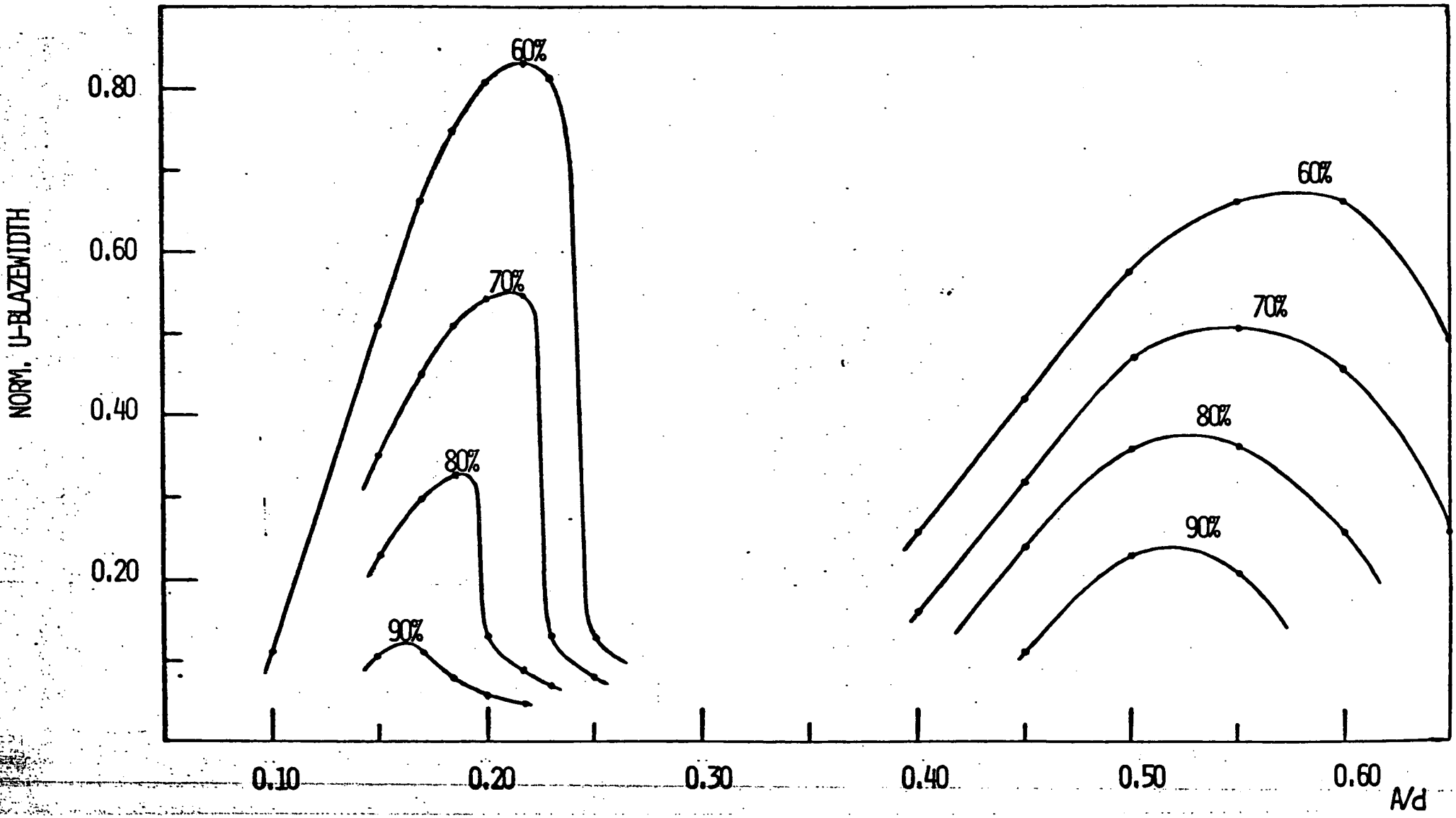


Figure (6.1) Normalized U-blazewidth curves as a function of profile amplitude for the sinusoidal grating in a Littrow mount. The curves show the order - 1 blazewidths for minimum efficiency levels of 60%, 70%, 80% and 90%.

is shown in figure (6.1). The blazewidth curves for P and S polarized radiation are given in reference (6.10).

The blazewidth curves shown here have two distinct peaks. The first peak (corresponding to shallow grooves) is associated with normalized wavelengths near 0.80 and angles of incidence close to 24^{o}. The second peak occurs for normalized wavelengths of about 1.20 and angles of incidence around 37^{o}. Thus if the two peaks are re-normalized to lie about the same central wavelength, the blazewidth for the upper peak will be multiplied by a factor of two-thirds. Using this weighting factor it is seen that the shallow-groove peak is the superior of the two for minimum efficiency levels of 60%, 70% and 80%.

In figure (6.2) the efficiency curves corresponding to the optimum 80% U-blazewidth (which occurs for A/d = 0.185) are shown for P, S and U polarizations. The important role that Wood anomalies play in the formation of a good blaze action for the sinusoidal profile is evident from these curves. (The Wood anomaly is indicated by the vertical bar at $\lambda/d = 0.66667$.)

6.3. ANOMALIES IN THE BLAZE REGION

The development of the P polarization anomaly evident in figure (6.2) is the most important factor influencing the strengthening of the U - blazewidth for shallow grooves. This anomaly is composed of two components, a Wood anomaly and a resonance anomaly. The resonance anomaly grows rapidly in strength once the profile amplitude A/d is greater than 0.10. For deeper profiles, the resonance anomaly separates from the Wood anomaly and moves towards longer wavelengths. Its shape is largely independent of profile amplitude after it separates from the

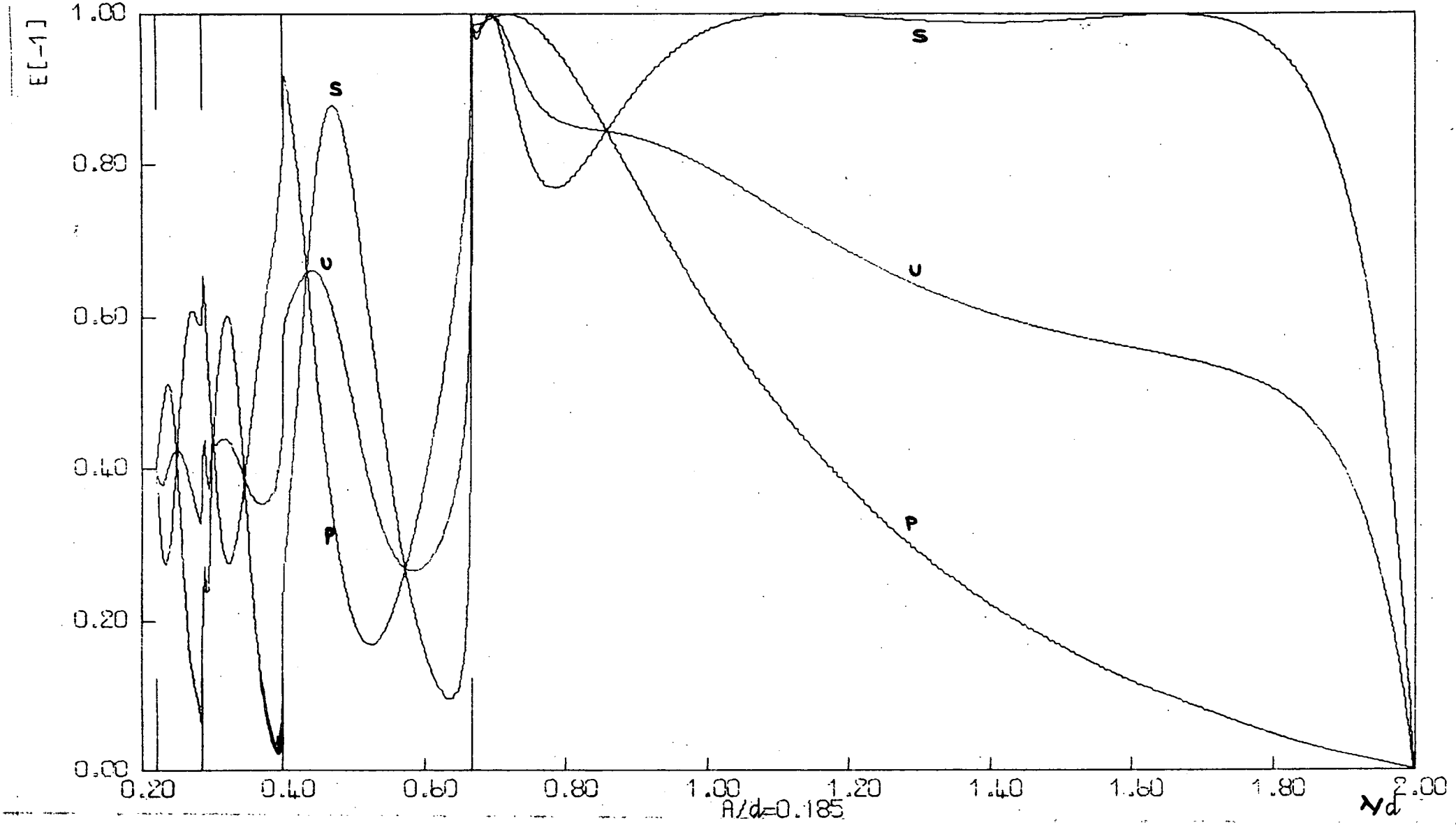


Figure (6.2) Order -1 efficiency curves for a sinusoidal grating having a normalized profile amplitude A/d = 0.185. The U polarization curve corresponds to the optimum blazewidth at the 80% efficiency level using the Littrow mount.

Wood anomaly.

The minimum near $\lambda/d = 0.80$ in the S polarization efficiency curve of figure (6.2) is caused by resonance in the evanescent orders -2 and +1. This minimum becomes more marked as the profile amplitude increases through $A/d = 0.20$; this causes the division of the blaze peak into two parts. It is this division which is responsible for the steep decline evident in the 60%, 70% and 80% blazewidth curves of figure (6.1).

In the amplitude region between the two U-blaze peaks the blazewidths remain very small. This is due not only to the action of the first S resonance but also to the emergence of a strong second S resonance near $A/d = 0.30$. The detailed behaviour of these resonances has already been discussed in Chapter 4.

In the region of profile amplitudes near the second blazewidth peak, the P polarization resonance peak in order -1 lies near an S polarization efficiency maximum. As the groove depth increases through $A/d = 0.50$ the development of an S polarization efficiency minimum causes the appearance of a downward trend in blazewidth values. This minimum is associated with resonance maxima in the intensities of the evanescent orders -2 and +1. Thus the same process which curtailed the growth of the first blaze peak causes the decay of the second blaze peak.

6.4. CONCLUSIONS

The efficiency curves for sinusoidal gratings in a Littrow mount show useful blaze properties in the order -1 only in the spectral region where two orders are propagating (i.e. for wavelengths greater than 0.67). In the wavelength regions where four or more orders propagate the unpolarized efficiency rarely rises above 60%.

The blazewidth curves of figure (6.1) show that the optimum profile amplitude varies slightly with the minimum efficiency levels. The 80% U-blazewidth curve defines an optimum value of A/d = 0.185. The optimum normalized blazewidth is 0.33. By comparison, the corresponding optimum value for a triangular profile grating having the same period is 0.32 (from Chapter 5). This comparability of the sinusoidal and triangular profile blazewidths is maintained for the other minimum efficiency levels.

In order to exploit the good blaze properties of optimum sinusoidal profiles, the groove depth must be precisely controlled, since the blazewidth peak for shallow profiles falls away rapidly on either side of the optimum.

The geometrical interpretation of the blaze action for triangular profile gratings cannot be applied to sinusoidal gratings. Instead the blaze properties of the latter are determined by the behaviour of the Wood anomalies and resonances for S and P polarization.

If the results discussed here are to be applied to spectral regions in which the reflectivity of the grating surface differs markedly from 100%, then discrepancies may occur. The recently published finite conductivity theory of Maystre (6.11) has shown that a peak efficiency of 100% for P polarized light at $\lambda = 0.75$ μm is lowered to 85% if the grating surface has the conductivity of evaporated aluminium. For such spectral regions the period-rescaling property mentioned above will no longer be valid. The close agreement in the form of Maystre's efficiency curves for P polarized light incident on infinitely conducting and aluminium surfaces, suggests that the optimum profile depth determined here will still be valid for sinusoidal gratings used in the visible spectrum.

REFERENCES

(6.1) McPhedran, R.C., Wilson, I.J., and Waterworth, M.D., Opt.Commun., in press.

(6.2) Brandes, R. and Curran, R., 1971, Appl.Optics, 10, 2101.

(6.3) Stroke, G.W., 1960, Rev.Opt., 39, 291.

(6.4) Bousquet, P., 1962, Rev.Opt., 41, 277.

(6.5) Blakely, J.M. and Olson, D.L., 1968, J.opt Soc.Amer., 39, 3476.

(6.6) Petit, R., 1966, Rev.Opt., 45, 249.

(6.7) Petit, R., and Maystre, D., 1972, Rev.de Phys.appliquée, 7, 427.

(6.8) Pavageau, J. and Bousquet, J., 1970, Opt.Acta , 17, 469.

(6.9) Nevière, M., and Cadilhac, M., 1971, Opt.Commun., 3, 379.

(6.10) Wilson, I.J., 1972, unpublished Honours Thesis (University of Tasmania, Physics Department).

(6.11) Maystre, D., 1972, Opt.Commun., 6, 50.

CHAPTER 7

PROFILE FORMATION IN HOLOGRAPHIC DIFFRACTION GRATINGS

This chapter is based on a paper by the author, Mr. I.J. Wilson and Dr. M.D. Waterworth (7.1), which has been accepted for publication by Optics and Laser Technology.

7.1. INTRODUCTION

Since 1967, much interest has been shown in the generation of spectroscopic diffraction gratings by holographic means. This method offers freedom from the restrictions and problems imposed by diamond wear, interferometric servo-control and the extensive high precision engineering encountered with classical ruling engines.

Holographic gratings also have the advantages of being virtually ghost-free and of having less stray light than classical gratings. High line densities are readily attainable and the size of the grating is only determined by the available laser power and the dimensions of the formation optics.

The profiles which are normally produced by ruling diamonds are composed of approximately linear segments, and traditionally take the form of two sides of a triangle, with the possible addition of unruled blank surface (or land). However, since holographic gratings are formed in an entirely different manner, the resultant groove profiles may be made to differ considerably from those produced by ruling engines.

We will discuss briefly the methods of producing holographic gratings, and then outline two methods of calculation of the profiles of such gratings. Of the two methods, only the second

yields profiles in agreement with all experimentally determined profiles. We will then consider the flexibility of arrangements by which holographic gratings can be made, and discuss the form of one such arrangement we have under development. The wide range of profiles which may be generated with such a system is also demonstrated.

7.2. PROFILE FORMATION

In this section we discuss the two current techniques of holographic grating production and the associated profile types. This leads to a model of profile formation, in photoresists, based on preferred chemical solution directions. A new production technique based on this model will enable a wide range of grating profiles to be made.

7.2.1. Current Techniques

Figure (7.1) illustrates two different ways, A and B, of forming holographic gratings. In both cases, two beams of coherent light are used to form an interference pattern in a thin photoresist layer set on a glass substrate. The intensity distribution of the interference pattern governs the rate of chemical solution of the photoresist during development. This gives rise to a periodic modulation of its thickness. The structure so formed can be used as a transmission grating, or it can be coated to form a reflection grating. Gratings having up to 3,600 lines/mm, and dimensions up to 16.5 cm x 32 cm have been produced in this way (7.2).

The arrangement A of figure (7.1) has been the one most frequently used in the past for the production of holographic gratings (7.3, 7.4). However, arrangement B, due to Sheridon (7.5), is also of

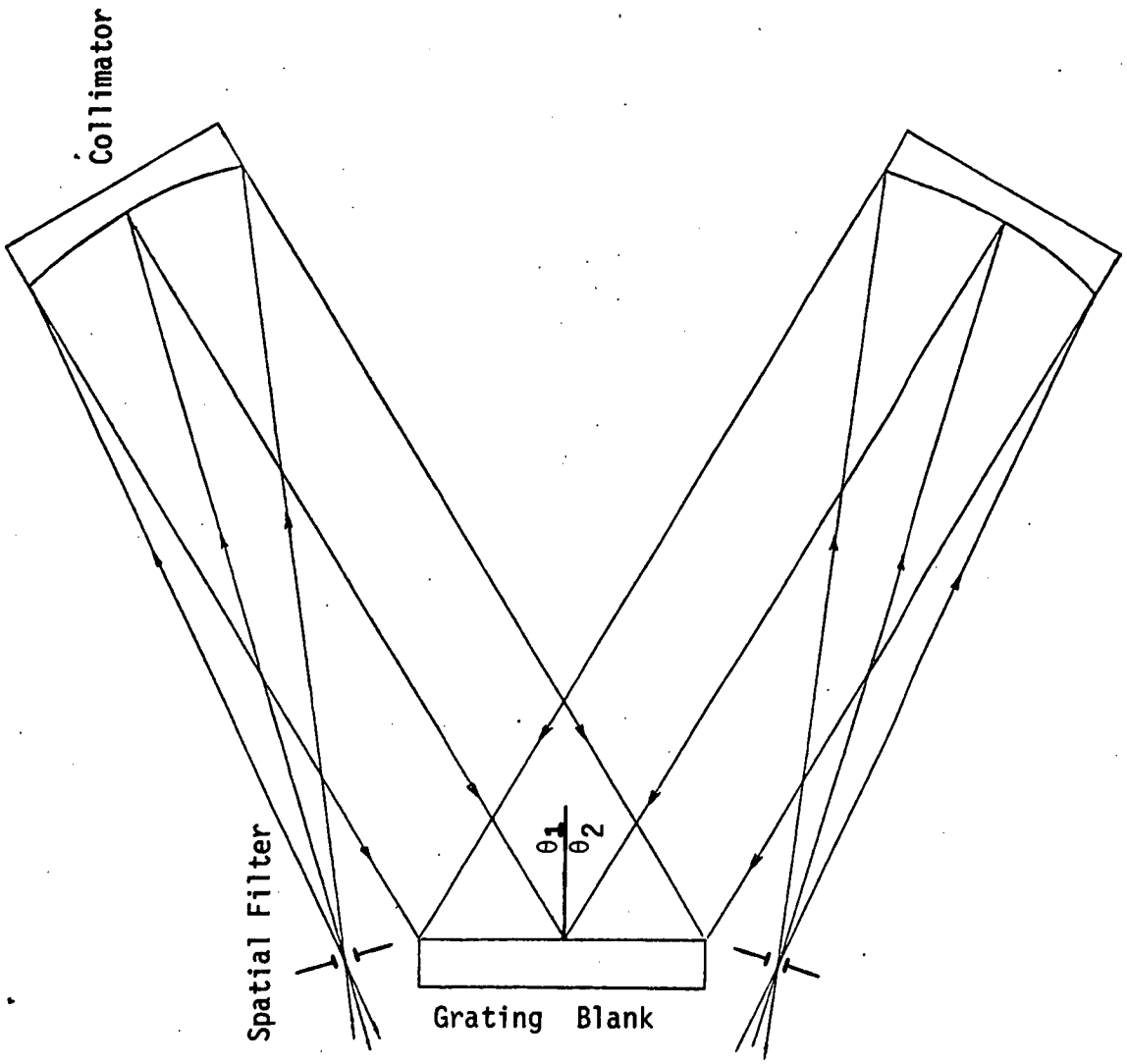


Figure (7.1). (Arrangement A) Elementary holographic grating formation. The interfering plane waves are incident on the front surface of the photoresist coated grating blank.

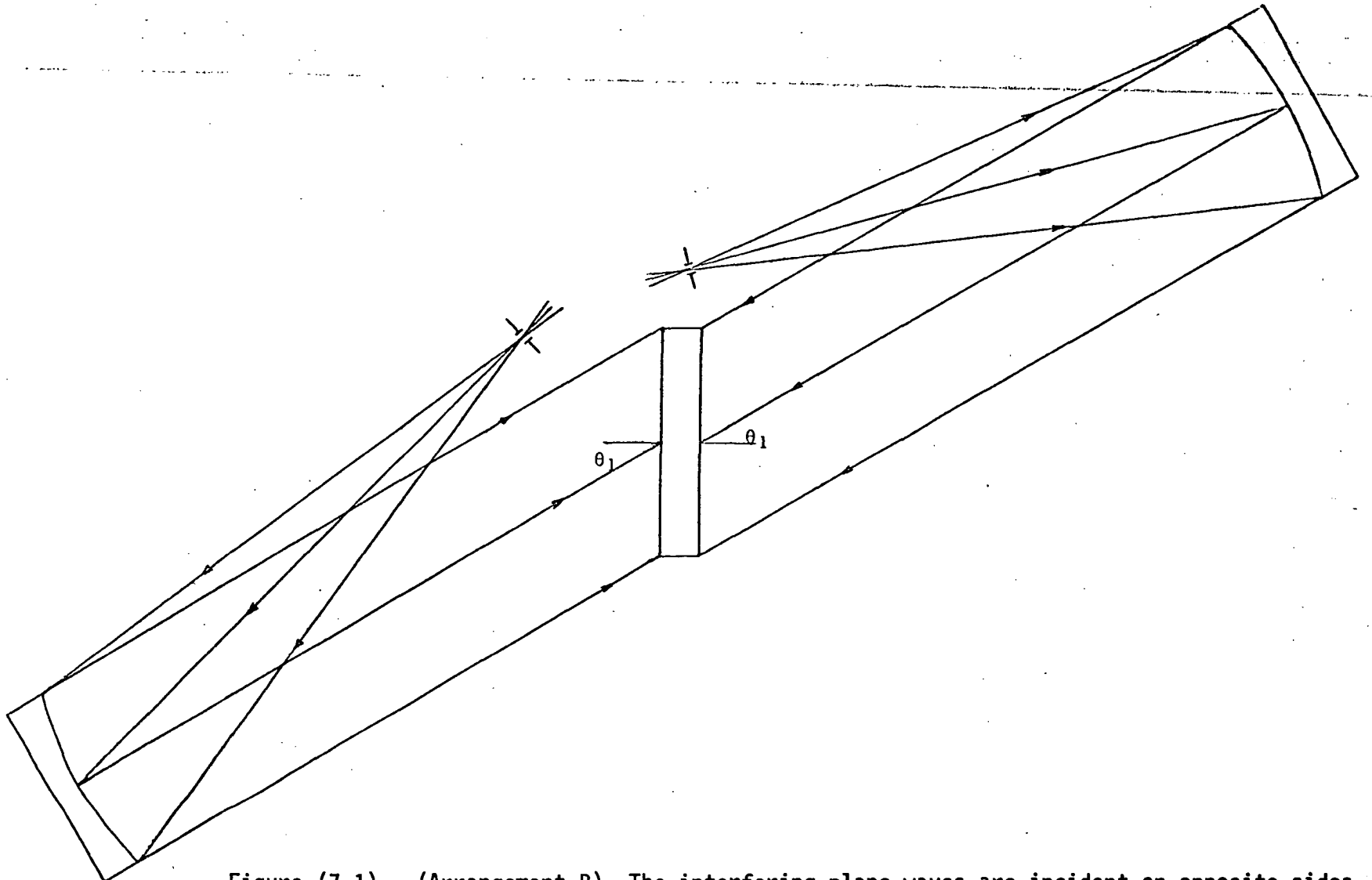


Figure (7.1). (Arrangement B) The interfering plane waves are incident on opposite sides of the photoresist coated grating blank.

interest, since it yields a quasi-triangular profile.

From Chapter 1, the exposure distribution corresponding to these two arrangements is of the form

$$E(x,y) = \tau\{1 - c \cos [\frac{2\pi x}{d} + \frac{2\pi n y}{\lambda} (\cos \theta'_1 - \cos \theta'_2)]\} \qquad (7.1)$$

where τ is the exposure time, c the fringe visibility, λ the wavelength of coherent light and n the refractive index of the photoresist. The Oy-axis is normal to the plane of the grating blank, while the Ox-axis lies in the plane of incidence of the two beams. The period, d, of the exposure is given by

$$d = \lambda/(\sin \theta_1 + \sin \theta_2) \qquad (7.2)$$

where the beam angles are θ_1 and θ_2 in air, and θ'_1 and θ'_2 in the photoresist. In the case of arrangement A, both θ_1 and θ_2 lie between 0 and $\pi/2$, but in arrangement B (the "standing wave" case), θ_2 lies between $\pi/2$ and π.

7.2.2. An Isotropic Model

We now consider an isotropic model of profile formation in photoresists. Suppose the profile ordinate corresponding to an exposure distribution E(x,y) at a time t in the development process is P(x,t). If we make the assumption that the rate of solution of photoresist at a particular point is isotropic and proportional to E at the point, then P satisfies the non-linear differential equation

$$\frac{\partial P(x,t)}{\partial t} = KE(x,y)\sqrt{1 + [\frac{\partial P(x,t)}{\partial x}]^2} \quad , \qquad (7.3)$$

where K is a constant determined by the photoresist solubility. This equation can be solved by the method applied to the problem of ray tracing in a non-homogeneous medium, in a way which has been described elsewhere (7.6).

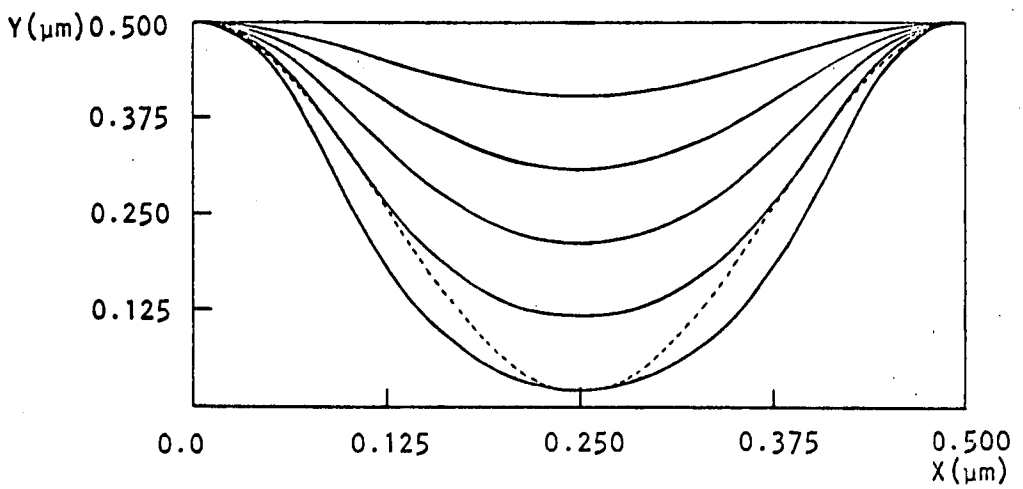


Figure (7.2). Isotropic - model profile curves.

The full lines delineate profiles calculated by the isotropic model of the solution process. As in all the following profile diagrams, the top straight line parallel to the Ox axis corresponds to the initial photoresist surface, and the five profile curves represent the situation at successive, equally spaced time intervals during the development process. The dashed line is a sinusoidal curve having the same depth as the final profile. Exposure data: $\theta_1 = \theta_2 = 29.2^{o}$, $\tau_1 = c_1 = 1.0$, linear positive photoresist, exposing wavelength $\lambda = 0.488\mu m$.

A profile obtained by the numerical integration of equation (7.3) is displayed in figure (7.2), together with a sinusoidal comparison curve having the same depth as the profile. Figure (7.2) corresponds to a symmetrical two beam exposure from the front of the grating blank ($\theta_1 = \theta_2$), so that the exposure function

$$E(x,y) = E(x) = \tau\{1 - c \cos \frac{(2\pi x)}{d}\} \qquad (7.4)$$

is a one-dimensional sinusoidal function.

Electronmicrographs obtained by Brandes and Curran (7.7), and reproduced in figure (1.10), show that the profiles formed in a linear photoresist using a one-dimensional, sinusoidal exposure function are themselves close to sinusoids. Figure (7.2), however, shows that the model based on isotropic photoresist solution predicts substantial deviations from sinusoidal form for such profiles. The model is therefore not suitable in this case.

7.2.3. Preferred Solution Direction Model.

If the process of photoresist solution is non-isotropic, the theoretical model must then distinguish some preferred direction. We choose this direction to be that of the lines of constant exposure in the photoresist, i.e. the directions of energy propagation in the medium. From equation (7.1), these are of the form

$$\frac{2\pi x}{d} + \frac{2\pi n y}{\lambda} (\cos \theta'_1 - \cos \theta'_2) = \text{constant} \qquad (7.5)$$

i.e. they are straight lines whose orientation is governed by the geometric configuration of the two interfering beams. If $\theta_1 = \theta_2$ then the lines of constant exposure are parallel to the Oy-axis.

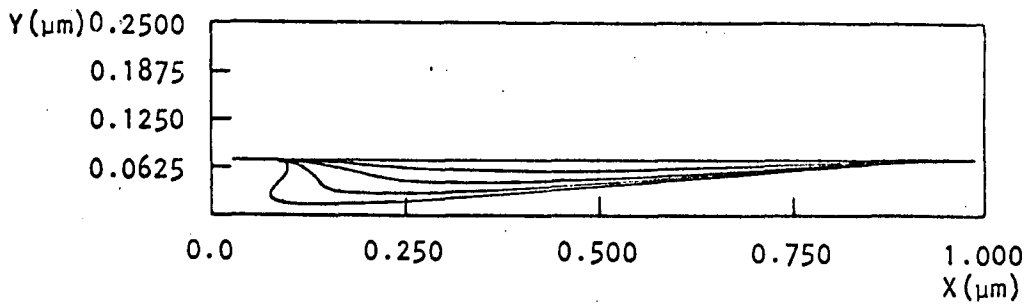


Figure (7.3). Quasi-triangular profiles.

The curves represent profiles obtained in a linear photoresist, for a single standing-wave exposure.

Exposure data: $\theta_1 = 180^o$— $\theta_2 = 14.1^o$, $\tau_1 = c_1 = 1.0$, linear positive photoresist with a solubility factor K = 0.005, $\lambda = 0.488\mu m$.

We will now assume that solution proceeds along lines of constant exposure, at a rate governed by the exposure function. In the case of a one-dimensional exposure ($\theta_1 = \theta_2$), solution will take place parallel to the Oy-axis, at a rate determined by E(x), i.e. the solution rate is independent of position along the solution track. Hence the profile function P(x,t) is directly proportional to E(x) and t. This means that the preferred solution direction model correctly predicts the formation of sinusoidal profiles with one dimensional exposures in linear photoresists (as in arrangement A of figure (7.1)).

If arrangement B of figure (7.1) is used, a resultant profile of the form illustrated in figure (7.3) is obtained. This shows the profile which is obtained when a linear positive photoresist is exposed by a standing wave interference pattern ($\theta_2 = \pi - \theta_1$). The computed profile is in good agreement with the electronmicrograph published by Sheridon (7.5), and reproduced in figure (1.11). As remarked above, this configuration of exposing beams provides an approximation to a shallow triangular profile.

This model can easily be adapted to the problem of calculating profiles in photoresists having non-linear exposure responses. This is achieved by calculating an effective exposure function, from the exposure function E(x,y), in a manner which takes into consideration the particular response function of the photoresist.

Figure (7.4) shows a profile curve calculated from a one dimensional exposure function, in the case of a logarithmic, negative photoresist. The response function was chosen on the basis of experimental data given in reference (7.2). The profile curve of figure (7.4) has a form in keeping with the electronmicrograph shown

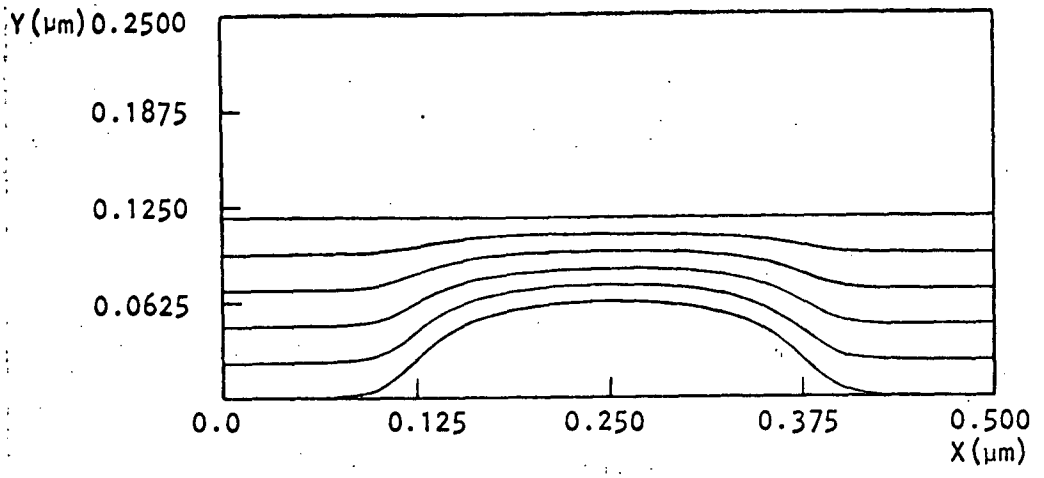


Figure (7.4). Profile curves for a logarithmic, negative photoresist.

Exposure data: $\theta_1 = \theta_2 = 29.2^o$, $\tau_1 = 30.0$, $c_1 = 1.0$, $K = 0.005$, $\lambda = 0.488\mu m$.

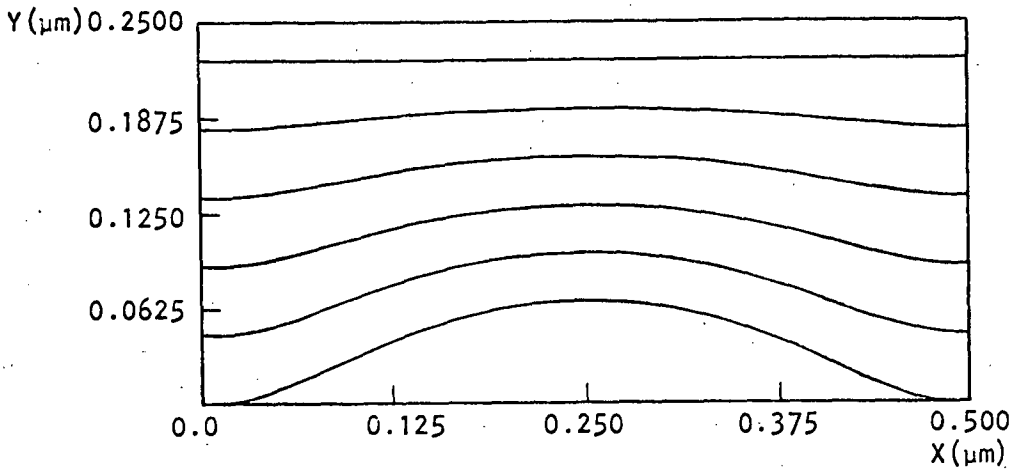


Figure (7.5). The effect of pre-exposure on the profile curves of figure (7.4).

Exposure data: $\theta_1 = \theta_2 = 29.2^o$, $\tau_1 = 50.0$, $c_1 = 0.28$, $K = 0.015$, $\lambda = 0.488\mu m$.

in figure (1.9) of a holographic grating (7.8) made by the firm of Jobin-Yvon in the logarithmic negative photoresist Resifax A. The flat regions of the profile correspond to exposure values below the sensitivity threshold of the photoresist. They can be removed if an appropriate pre-exposure is made with a single beam of coherent light, before the two beam interference pattern is recorded. The pre-exposure can be taken into account by using a fringe contrast factor, c, smaller than unity in the theoretical profile calculation. Figure (7.5) shows the effect of a pre-exposure on the profile curve of figure (7.4). The land has been removed from the profile at the cost of an increase in the exposure time required (because of the decrease in fringe contrast).

The preferred solution direction model of photoresist development is thus seen to yield results in keeping with all experimental evidence now available on holographic diffraction grating profile forms.

7.2.4. Multiple Exposures.

The two-beam configurations of figure (7.1) yield only a limited range of forms of the exposure function, and thus do not give great freedom in the choice of the resulting profile forms. We will now consider the possibility of generalizing these arrangements to give a wider range of exposure functions.

A multiple-beam exposure can always be considered to result from the superposition of single two-beam exposures. Multiple two-beam exposures using only variations of the phase difference between the interfering beams are equivalent to a single two-beam exposure with appropriate amplitude and phase. Thus, we must utilize different angular configurations of the interfering beams, during any multiple exposure, in order to realize a resultant exposure function differing

from those obtained with current methods.

Such a multiple two-beam exposure is characterized by an expression of the form

$$E(x,y) = \sum_{m=1}^{M} \tau_m \{1 - c_m \cos [\frac{2\pi x}{\lambda} (\sin \theta_{1,m} + \sin \theta_{2,m}) + \frac{2\pi ny}{\lambda} (\cos \theta'_{1,m} - \cos \theta'_{2,m}) + \phi_m]\} \qquad (7.6)$$

where M is the number of two-beam exposures, each of which is specified by a partial exposure time τ_m, a phase factor ϕ_m, a fringe contrast factor c_m and beam angles $\theta_{1,m}$ and $\theta_{2,m}$.

With the exposure function in this general form, there may not exist lines of constant total exposure through all points in the photoresist. Thus, solution paths must be along lines of constant partial exposure. At a particular profile point, solution continues along that direction which will result in maximum profile movement in a given increment of time, i.e. along that line of constant partial exposure which lies closest to the profile normal at the point in question. The rate of solution along the preferred direction is controlled by the total photoresist exposure.

7.3. PROPOSED EXPERIMENTAL TECHNIQUES

In this section we consider the tolerances on the intensity variation, phase stability and period control of each exposure that must be satisfied in order to produce a given profile using the multiple exposure method. Optical techniques capable of satisfying these requirements are then briefly discussed.

The computer generated profiles indicate that the intensity

distribution across the aperture of the grating blank due to each collimator should be uniform to within ±5%. Since the intensity distribution across the wavefronts propagating from the collimators in the basic arrangement will be a Gaussian intensity distribution, this criterion may not be satisfied. Approximately uniform illumination may be achieved if the interfering beams are expanded sufficiently or if the laser light is passed through a Gaussian absorption filter.

The theory indicates that the relative phase difference between any two exposure distributions must be controlled to within $\pm \pi/8$. This can be achieved by using one of the feedback systems for fringe stabilization now used in holography (7.9). A feedback signal may be obtained by monitoring a coarse fringe pattern produced by interfering part of the light from each arm of the grating production interferometer. This signal then controls a phase shifting element affecting one of the beams (see figure (7.6)).

The most difficult problem in producing general profiles using the multiple exposure system is one of controlling the period of the exposure distribution. From a consideration of the effect of a change in the relative phase in the computer generated profiles, a relative period tolerance of approximately $\frac{\delta d}{d} = \pm \frac{1}{16N}$ for the smallest exposure period, d, is necessary. Here N is the number of lines from the centre of the blank to its edge.

To achieve the angle setting tolerance required to obtain period control of this accuracy, consider equation (7.2), which for symmetric exposures can be written as:

$$\sin \theta = \lambda N/2h \qquad (7.7)$$

where 2h is the blank aperture. Differentiating with respect to N in

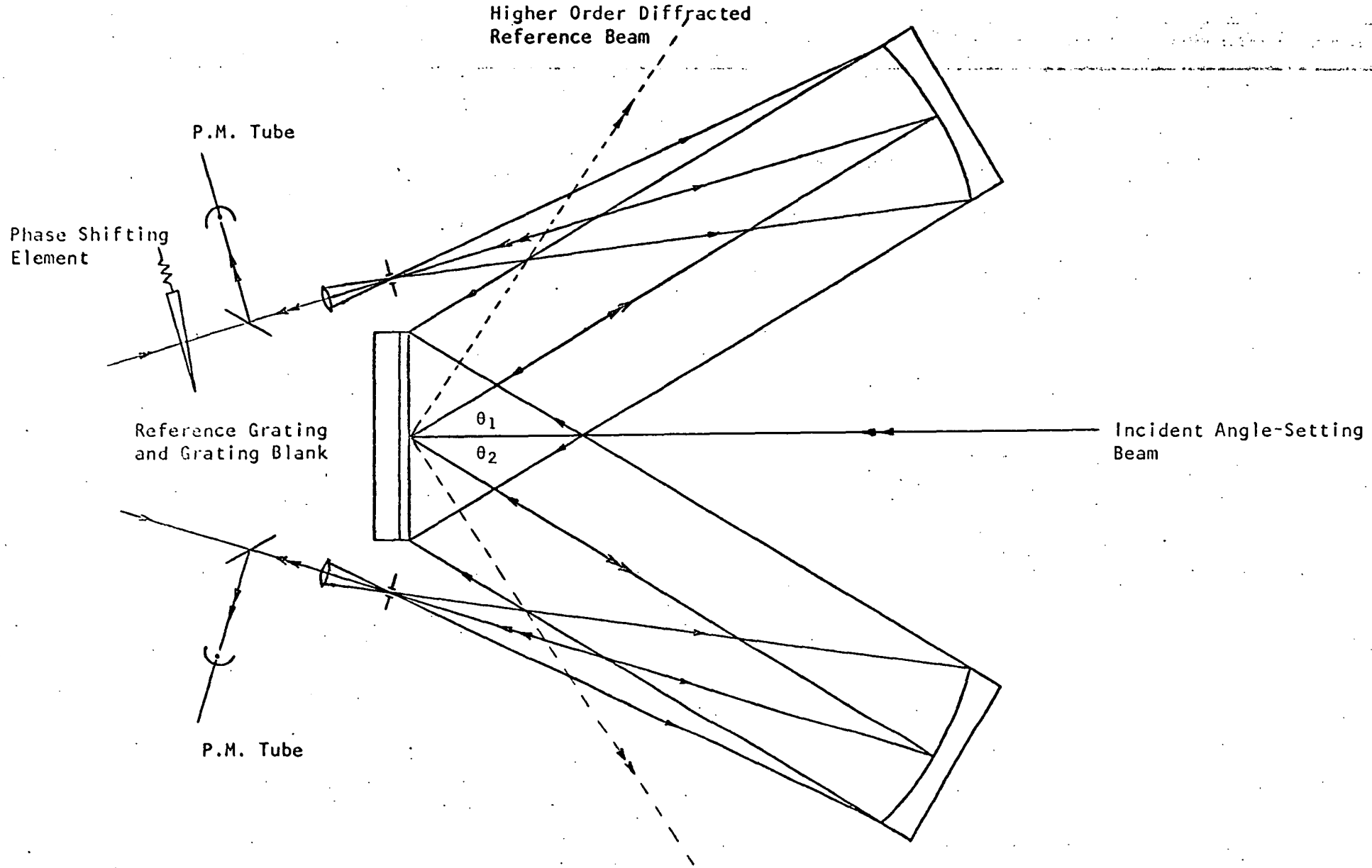


Figure (7.6). Multiple Exposure Holographic Grating Formation

equation (7.7), the angle tolerance, $\delta\theta$, is

$$\delta\theta = \lambda\delta N/(2\ h\ \cos\ \theta)\ . \qquad (7.8)$$

A tolerable phase error of $\pm\ \pi/8$ corresponds to a line error $\delta N = \pm\ 0.0625$. Thus for a 10 cm × 10 cm grating, h = 5 cm and $\theta = 29.2^{o}$, so that $\delta\theta = \pm\ 0.35\ \mu R$.

To define the angles $\theta_{1,m}$ and $\theta_{2,m}$ to within the tolerance, $\delta\theta$, it is proposed to add a reference grating to the basic arrangement of figure (7.1)A, in the plane of the grating blank. It will be illuminated with collimated laser light of the exposing wavelength, λ, which is diffracted into the various orders, given by the grating equation:

$$\sin\ \theta_d - \sin\ \theta_i = m\lambda/d_R\ . \qquad (7.9)$$

In equation (7.9), θ_i is the angle of incidence relative to the grating normal, θ_d is the diffracted angle, m is the diffracted order number, and d_R is the period of the reference grating.

The diffracted angles of orders $\pm$ m will be given by:

$$\sin\ \overset{(+m)}{\theta_d} = \sin\ \theta_i + (m\lambda/d_R),\ \overset{(+m)}{\theta_d} > 0,$$

and

$$\sin\ \overset{(-m)}{\theta_d} = \sin\ \theta_i - (m\lambda/d_R),\ \overset{(-m)}{\theta_d} < 0\ .$$

Then

$$\sin\ \overset{(+m)}{\theta_d} - \sin\ \overset{(-m)}{\theta_d} = 2m\lambda/d_R\ . \qquad (7.10)$$

But $\overset{(-m)}{\theta_d} < 0$, so equation (7.10) gives

$$\sin\ \overset{(+m)}{\theta_d} + \sin\ |\ \overset{(-m)}{\theta_d}\ | = 2m\lambda/d_R\ . \qquad (7.11)$$

Now $$\sin \theta_d^{(+m)} = \sin \theta_1 \text{ and } \sin \left| \theta_d^{(-m)} \right| = \sin \theta_2$$

in the grating formation equation (equation (7.2)), hence

$$d = d_R/2m \ . \qquad (7.12)$$

Thus, the diffracted beams from the reference grating precisely define the angles $\theta_{1,m}$ and $\theta_{2,m}$ required to produce a holographic grating having exposure periods $d_R/2m$.

The spatial filter and collimator of each arm of this interferometer will be fixed relative to each other so that by rotating each arm about a centre under the grating blank, the angles $\theta_{1,m}$ and $\theta_{2,m}$ can be varied. Using the light diffracted by the reference grating, the collimator mirrors can be set up so that the light reflected towards their focii passes through the spatial filter pin-holes. Part of this beam will be sensed by a photomultiplier tube (see figure (7.6)). Each arm of the interferometer will then be rotated until maximum photomultiplier tube current is obtained for the appropriate reference order.

The resolution of this angle setting technique depends on the spot size at the focii of the collimators and the spatial filter pin-hole diameter. To produce a 10 cm × 10 cm grating of period 0.5 μm, the angle setting tolerance $\delta\theta$ is ± 0.35 μR. For 2m focal length collimators, this corresponds to a transverse shift of focus of ± 1.40 μm. If the spatial filter pin-hole diameter is 5 μm, the peak intensity of the light propagating to the photomultiplier tube may be sensed with sufficient accuracy to enable the desired angle setting tolerance to be realized.

7.4. EXAMPLES OF MULTIPLE EXPOSURE PROFILES

The exposure functions which correspond to the interference patterns used to form holographic gratings are smoothly varying. Hence a seemingly difficult problem would be to generate good approximations to profiles composed of straight line segments by superimposing a small number of sinusoidal or log-sinusoidal exposure functions. The theoretical profiles in figures (7.7, 7.8, 7.9) show two-term approximations to triangular and lamellar groove forms. The quasi-triangular profiles do not have perfectly straight facets, but the deviations from linearity are not large. The holographic approximation to the lamellar profile offers the possibility of constructing gratings for the visible region having the good blaze properties measured at millimetre wavelengths by Wirgin and Deleuil (7.10).

The triangular forms of figures (7.7) and (7.8) have been produced using the exposure configuration A of figure (7.1). Quasi-triangular forms can also be produced using the exposure configuration B of figure (7.1). However, the use of two exposures affords a greater degree of control of the profile blaze angle and depth than is obtainable with a single standing wave exposure. In both cases, the groove angles and profile depth change continuously during development, but the ranges of variation are much larger for the former method of formation. Thus, it is only necessary to vary the development time of the doubly-exposed photoresist in order to shift the blaze-wavelength of the holographic grating through a wide spectral region. Consequently the double exposure configuration seems capable of providing a wider range of values of the ratio of blaze wavelength to period, than does the configuration of figure (7.1)B.

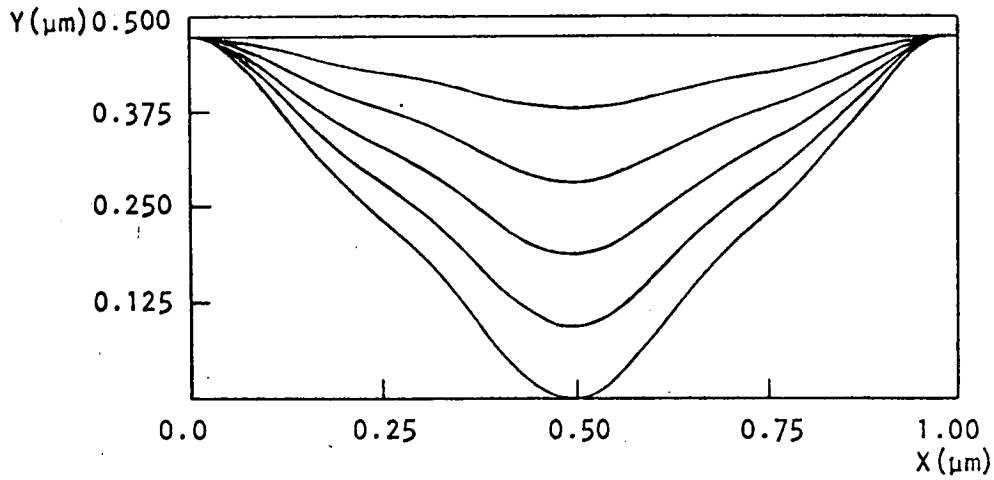


Figure (7.7). Approximately triangular, symmetrical profiles produced by two exposures.

Exposure data: linear positive photoresist , K = 0.005, λ = 0.488μm.

Exposure 1: d_1 = 0.3333μm, τ_1 = 0.10, ϕ_1 = 0.0, c_1 = 1.0, $\theta_{1,1} = \theta_{2,1} = 47.10^{o}$

Exposure 2: d_2 = 1.0μm, τ_2 = 0.90, ϕ_2 = 0.0, c_2= 1.0, $\theta_{1,2} = \theta_{2,2} = 14.12^{o}$.

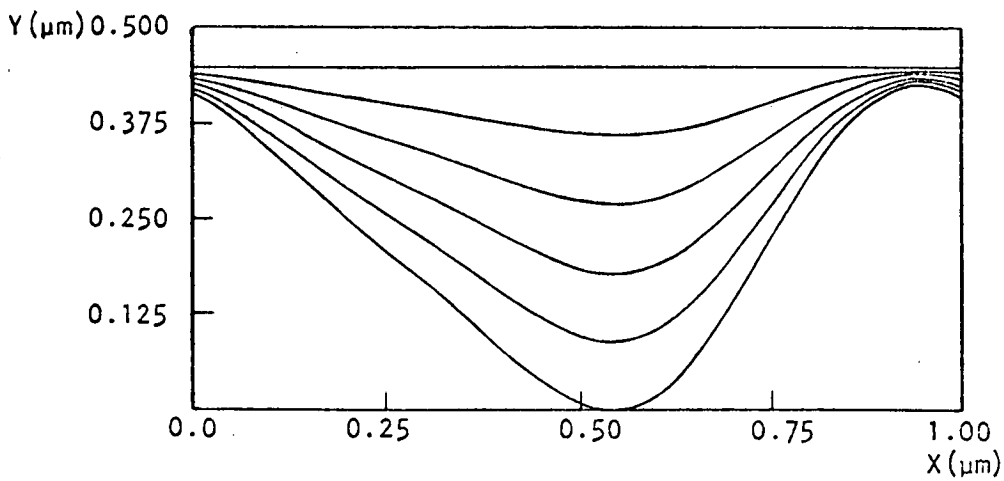


Figure (7.8). A two-exposure approximation to an asymmetric triangular profile.

Exposure data: linear positive photoresist, K = 0.005, λ = 0.488μm.

Exposure 1: d_1 = 0.5 μm, τ_1 = 0.15, $\phi_1 = \pi/2$, c_1 = 1.0, $\theta_{1,1} = \theta_{2,1} = 29.20^{o}$.

Exposure 2: d_2 = 1.0 μm, τ_2 = 0.85, ϕ_2 = 0, c_2 = 1.0, $\theta_{1,2} = \theta_{2,2} = 14.12^{o}$.

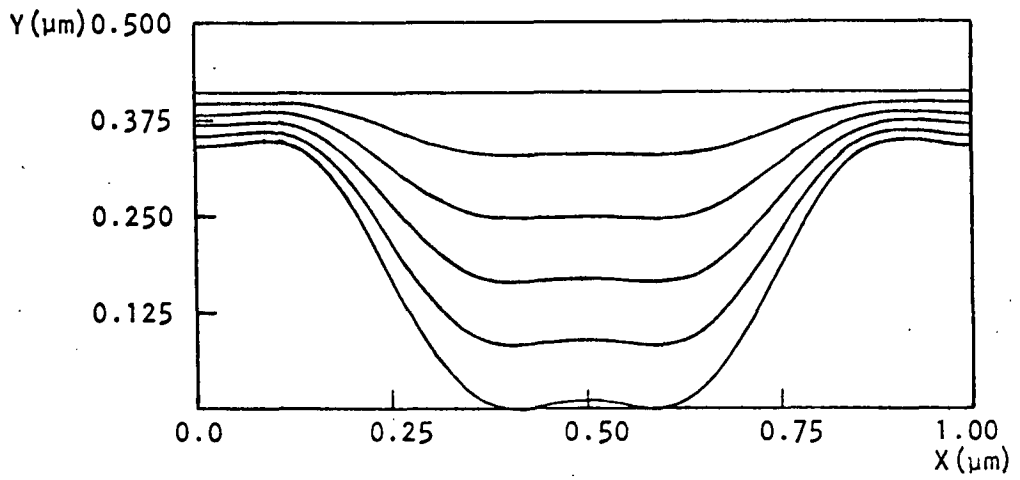


Figure (7.9). A two-exposure approximation to a rectangular profile.

Exposure data: linear positive photoresist, K = 0.005, λ = 0.488μm.

Exposure 1: d_1 = 0.3333μm, τ_1 = 0.15, ϕ_1 = π, c_1 = 1.0, $\theta_{1,1} = \theta_{2,1} = 47.10^{o}$.

Exposure 2: d_2 = 0.50μm, τ_2 = 0.85, ϕ_2 = 0, c_2 = 1.0, $\theta_{1,2} = \theta_{2,2} = 29.20^{o}$.

The final two profile curves, figures (7.10, 7.11), illustrate two groove shapes which differ in form from any shapes which have been investigated before, either experimentally or theoretically. Neither seems capable of generation by other than holographic means. The second profile has two quasi-triangular elements in each period, having different blaze angles and depths. If each of these elements has a blaze centred on a different wavelength, then by manipulation of the two blaze angles, high diffraction efficiencies may well be achieved over a wider wavelength range than can be obtained with a conventional triangular profile.

7.5. CONCLUSIONS

It has been shown that a theoretical model based on a preferred direction of photoresist solution can yield profile forms of holographic gratings in good agreement with all experimentally determined forms. A model of profile formation based on isotropic solution is untenable from this viewpoint.

The form of an experimental apparatus which can provide multiple exposures having accurately known periods and phases has been suggested. The wide range of profile forms which could be generated using such an arrangement has been demonstrated.

Once freedom of choice of profile form has been achieved, it may well prove possible to greatly improve the spectral performance of diffraction gratings. Theoretical investigations of the diffraction properties of a wide class of profile forms will be necessary, in order to facilitate blaze optimization of multiple-exposure holographic gratings. In the two preceding chapters, studies

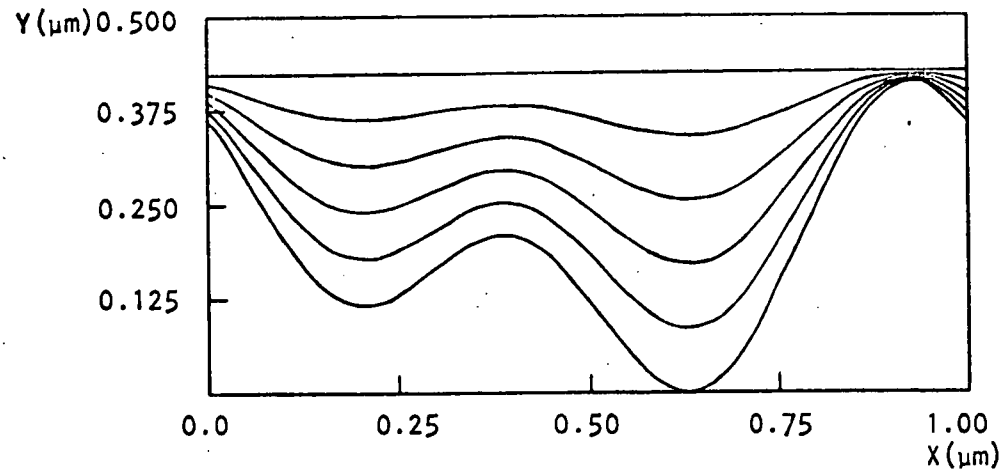


Figure (7.10). A new profile form produced by two exposures.

Exposure data: linear positive photoresist, K = 0.005, λ = 0.488μm.

Exposure 1: d_1 = 0.50μm, τ_1 = 0.50, ϕ_1 = $\pi/3$, c_1 = 1.0, $\theta_{1,1} = \theta_{2,1} = 29.20^{o}$.

Exposure 2: d_2 = 1.00μm, τ_2 = 0.50, ϕ_2 = 0.0, c_2 = 1.0, $\theta_{1,2} = \theta_{2,2} = 14.12^{o}$.

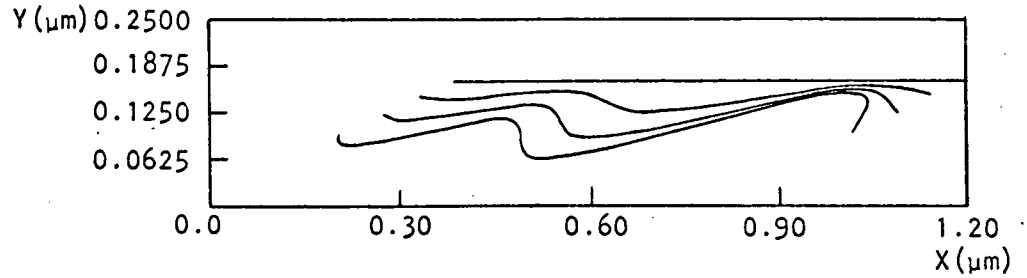


Figure (7.11). "Double-blazed" triangular profiles. The difference in blaze angle is approximately 2^{o}.

Exposure data: linear positive photoresist, K = 0.005, λ = 0.488μm.

Exposure 1: d_1 = 0.50μm, τ_1 = 0.70, ϕ_1 = 0.0, c_1 = 1.0, $\theta_{1,1} = \pi - \theta_{2,1} = 29.20^{o}$.

Exposure 2: d_2 = 1.00μm, τ_2 = 0.30, ϕ_2 = 0.0, c_2 = 1.0, $\theta_{1,2} = \pi - \theta_{2,2} = 14.12^{o}$.

of the blaze properties of the triangular and sinusoidal profiles have been presented. More general profile forms, including the interesting "double blaze" structure of figure (7.11), have yet to be considered in detail.

In addition to the improvements in blaze properties which can be achieved through suitable control of grating profile shape, further improvements may be obtained through the use of appropriately-shaped dielectric coatings placed on top of the grooves. M. Nevière, M. Cadilhac and R. Petit (7.11) have obtained encouraging results from a theoretical study of dielectric-coated, triangular profile gratings.

REFERENCES

(7.1) McPhedran, R.C., Wilson, I.J., and Waterworth, M.D., Optics and Laser Technology, in press.

(7.2) Jobin-Yvon, Company Publication.

(7.3) Labeyrie, A. and Flamand, J., 1969, Opt.Commun., 1, 5.

(7.4) Rudolph, D. and Schmahl, G., 1970, Optik, 30, 475.

(7.5) Sheridon, N.K., 1968, Appl.Phys.Letters, 12, 316.

(7.6) McPhedran, R. and Waterworth, M., 1971, Eighth Australian Spectroscopy Conference.

(7.7) Brandes, R. and Curran, R., 1971, Appl.Optics, 10, 2101.

(7.8) Maystre, D. and Petit, R., 1970, Opt.Commun., 2, 309.

(7.9) Neumann, D. and Rose, H., 1967, Appl.Optics, 6, 1097.

(7.10) Wirgin, A. and Deleuil, R., 1969, J.opt.Soc.Amer., 59, 1348.

(7.11) Nevière, M., Cadilhac, M., and Petit, R., 1972, Opt.Commun., 6, 34.

CHAPTER 8.

EFFICIENCY CURVES AND PROFILE VARIATION.

The theoretical efficiency curves shown in this chapter are based on computations carried out by the author in conjunction with Mr. I.J. Wilson. The method used for the calculation of grating efficiencies has already been described in Chapter 2 and Appendix I.

8.1. INTRODUCTION

In Chapter 2, it was mentioned that the quality of the agreement to be expected between the predictions of a rigorous, infinite conductivity theory and the behaviour of optical gratings was still undetermined. In this chapter, this problem will be investigated, both for gratings having a triangular profile and for gratings having the near-sinusoidal profiles which can be produced by holographic techniques. For the latter profile class, the important influence which profile distortion exerts on the form of grating anomalies will be demonstrated. It will be shown that this influence can be exploited in such a way as to improve the spectral performance of holographic gratings by at least 70%. The effects of the presence of unruled surface (or land) on the efficiency curve of a triangular profile grating will also be investigated.

8.2. TRIANGULAR PROFILE GRATINGS.

The efficiency curves shown in figures (8.1), (8.3) and (8.5) were communicated to the author by E.G. Loewen (8.1). The relative efficiency values were measured in a spectrometer having an 8^{o} included angle between incident and diffracted beams. The several sets of curves

on each graph correspond to gratings having differing line densities but (nominally) the same profile angles. Since efficiency is here plotted as a function of the Littrow diffraction angle (which is equal to the magnitude of the actual diffraction angle in order -1, less 4^{o}), the curves for the gratings of differing line densities would be identical, if the gratings had in fact the same profile shapes. The important differences between some corresponding curves make clear the practical difficulty of holding a ruled grating profile to a specified form.

For each of the three blaze angles, theoretical efficiency curves were computed for triangular profile gratings, used with the angular deviation in order -1 equal to -8.0^{o}. The groove apex angle was varied in 10^{o} steps until the best agreement between theoretical and experimental curves was obtained. Figures (8.2), (8.4) and (8.6) show the theoretical counterparts obtained in this way for the experimental curves of figures (8.1), (8.3) and (8.5).

Experimental efficiency curves are shown in figure (8.1) for gratings having blaze angles of 17.45^{o} and line densities of 600 and 1200/mm. The two P polarization curves are in quite close agreement, but the deviations between the S polarization curves are more significant. The S Wood anomalies associated with the passing-off of orders +1, -2 and +2 are much more pronounced for the 600 line/mm grating than for the 1200 line/mm grating.

The theoretical curves of figure (8.2) correspond to a grating having a blaze angle of 17.45^{o} and a groove apex angle of 100^{o}. For P polarization, theory and experiment are in good quantitative agreement everywhere except in the region just before the experimental blaze peak. The theoretical peak is about 2% higher than the measured

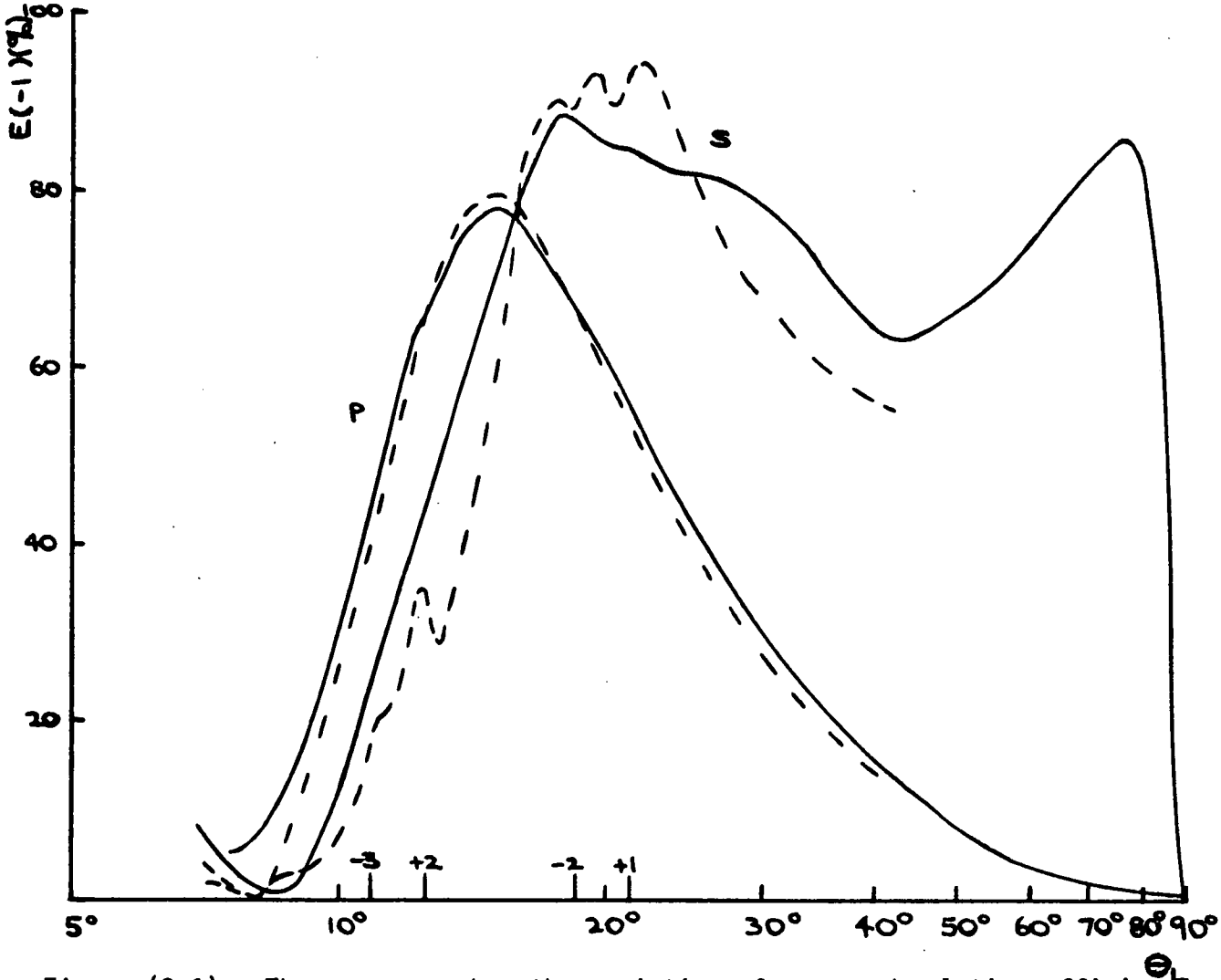


Figure (8.1). These curves show the variation of measured relative efficiency values in P and S polarized light with the Littrow diffraction angle, for two gratings each having a nominal blaze angle of 17.45^{0}, used with an angular deviation of -8.0^{0} in order -1. The solid lines correspond to the grating having 1200 grooves/mm, while the dashed lines refer to the grating having 600 grooves/mm. As in the following five figures, Rayleigh angles and the relevant grazing orders are marked on the abscissa axis. (After E.G. Loewen.)

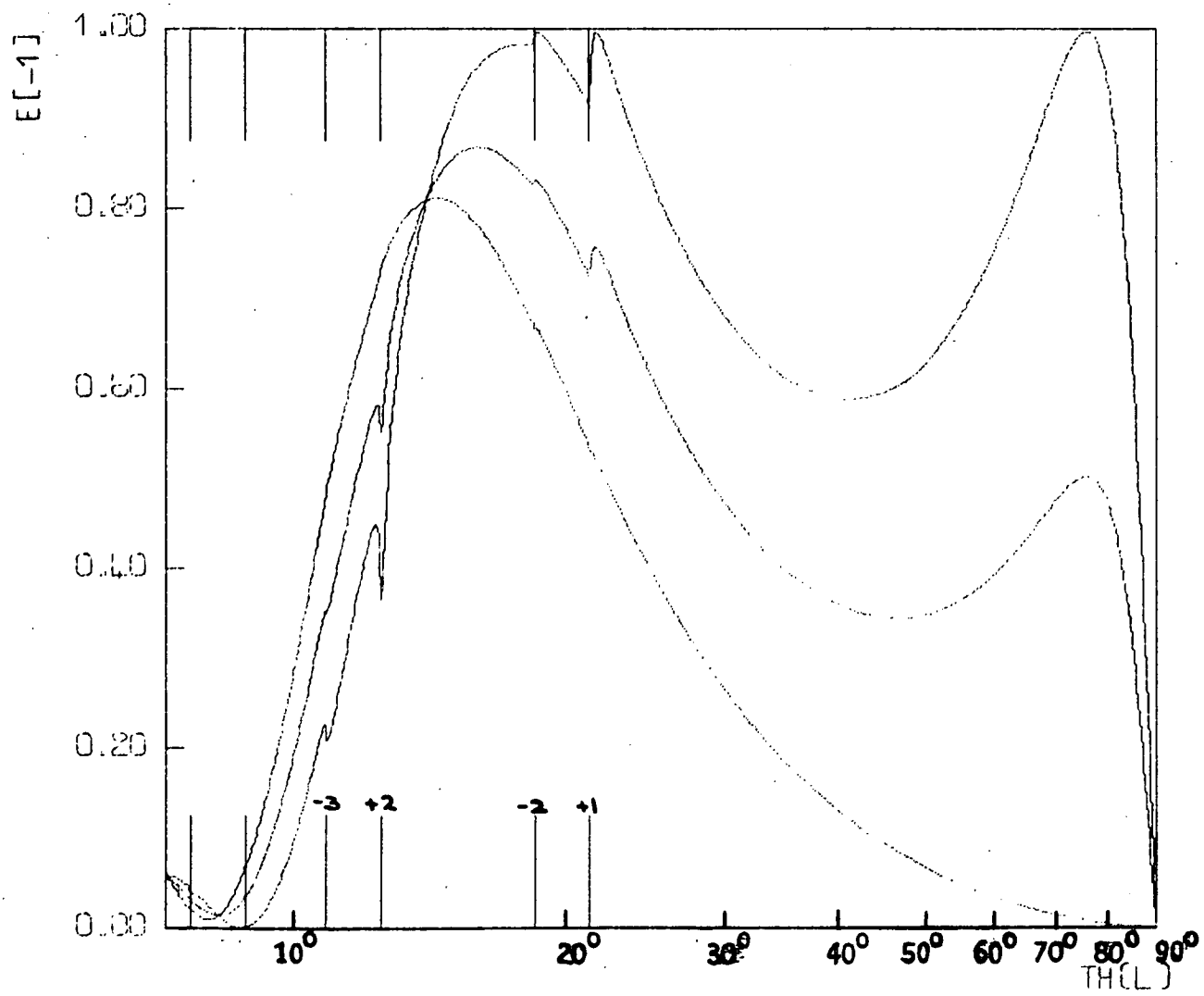


Figure (8.2). Theoretical efficiency curves are given for P, S and U polarized light, for a triangular profile grating having a blaze angle of 17.45° and a groove apex angle of 100°, used with an angular deviation of -8.0° in order -1. The abscissa is the Littrow diffraction angle, which is plotted on a logarithmic scale.

peak, and corresponds to a slightly lower diffraction angle. For S polarization, the theoretical curve is in quantitative agreement with that for the 600 line/mm grating in the angular region where only two orders propagate (i.e., above the Rayleigh angle for order +1). Both experimental curves drop away more sharply from their blaze peaks than does the theoretical curve. The angular discrepancy between the three is at most 2.5^{o}. The theory correctly predicts the form of all the anomalies given by the 600 line/mm grating.

Experimental efficiency curves are shown in figure (8.3) for gratings having blaze angles of 10.5^{o} and line densities of 1200 and 1500/mm. Despite the fact that the two gratings have comparable periods, their efficiency curves are in some regions quite dissimilar. For P polarization, the two blaze peaks differ in strength by 2%, and in angular position by 1.5^{o}. For S polarization, the two blaze peaks have quite different strengths and angular widths. The 1500 line/mm grating has three anomalies in the blaze region, one of which is quite strong, whereas the 1200 line/mm grating apparently forms no anomalies in this region. Both gratings give pronounced anomalous maxima after the passing off of orders -2 and +1.

The theoretical curves of figure (8.4) correspond to a grating having a blaze angle of 10.5^{o}, and a groove apex angle of 120^{o}. For P polarization, there is good quantitative agreement between the theoretical efficiencies and the experimental values for the 1500 line/mm grating. The two blaze peaks are located at the same angular position, and differ in strength by only 3%. For S polarization, the agreement between theory and experiment is less satisfactory. The theoretical curve correctly predicts the forms of all the anomalous features given by the grating having 1500 lines/mm. However, the experimental

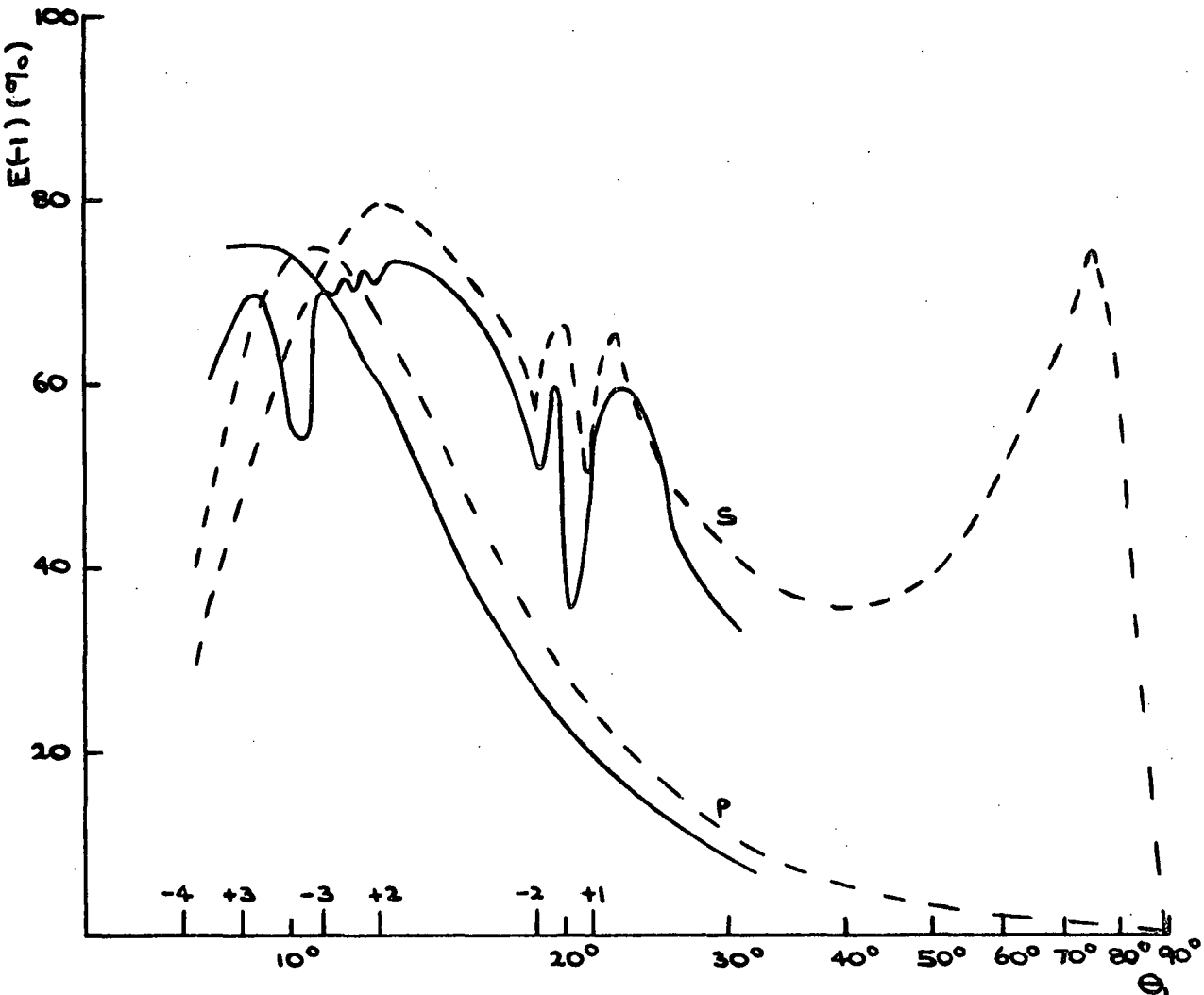


Figure (8.3). These curves show the variation of measured relative efficiency values in P and S polarized light with the Littrow diffraction angle, for two gratings each having a nominal blaze angle of 10.5^{0}, used with an angular deviation of -8.0^{0} in order -1. The solid lines correspond to the grating having 1500 grooves/mm, while the dashed lines refer to the grating having 1200 grooves/mm. (After E.G. Loewen.)

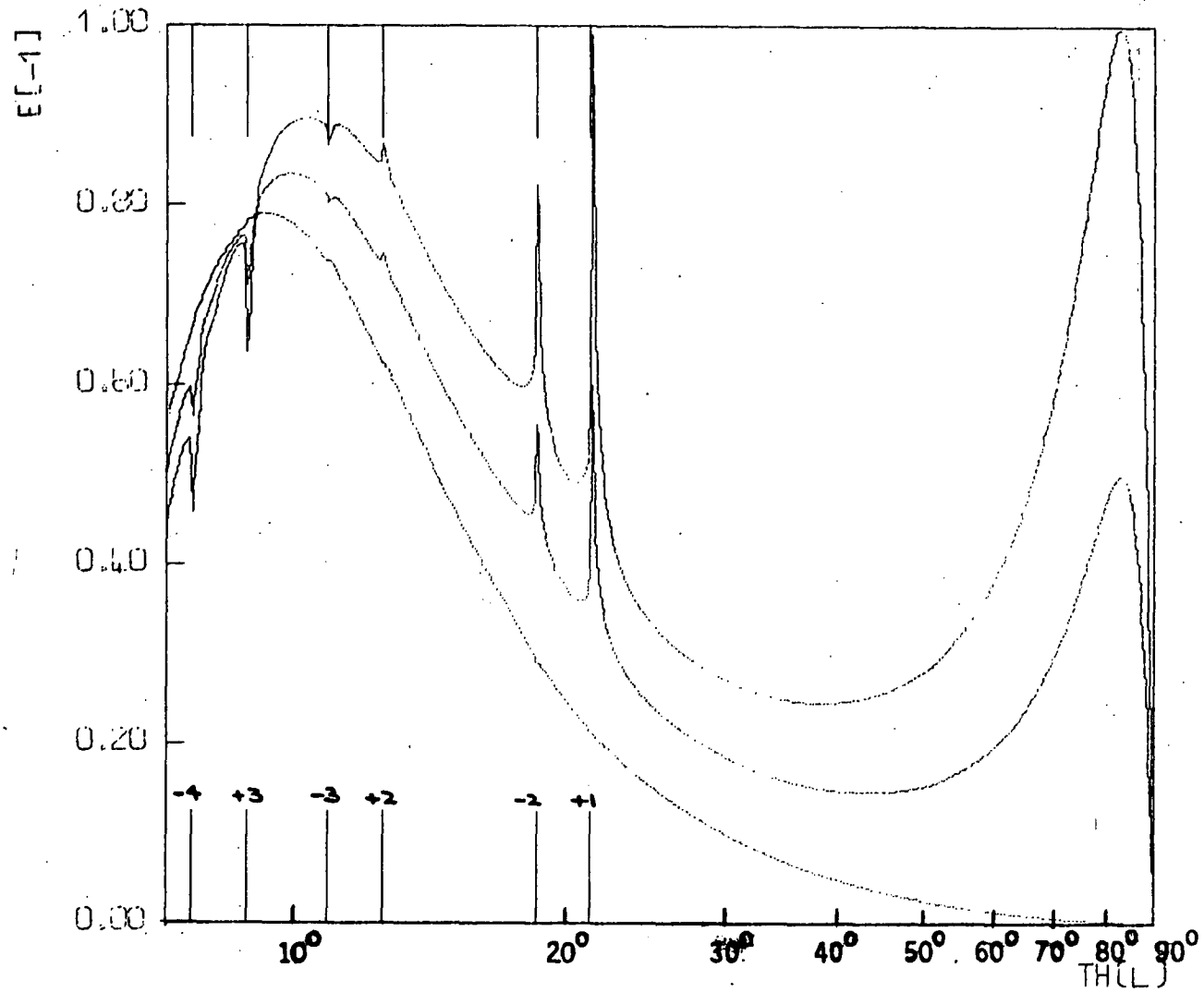


Figure (8.4). Theoretical efficiency curves are given for P, S and U polarized light, for a triangular profile grating having a blaze angle of 10.5^{0} and a groove apex angle of 120^{0}, used with an angular deviation of -8.0^{0} in order -1. The abscissa is the Littrow diffraction angle, which is plotted on a logarithmic scale.

anomalies tend to extend over wider angular ranges than do their theoretical counterparts. This is particularly evident for the minimum associated with the passing off of order +3, and for the maxima corresponding to resonances in the evanescent orders -2 and +1. (These angular discrepancies may in part be due to a lack of resolution in the experimental efficiency tester (8.1). We would also intuitively expect a broadening of the resonance features because of the influence of the finite conductivity of the experimental grating's surface.) The theoretical value of the peak efficiency for S polarization is 10% greater than the largest experimental efficiency.

Three sets of experimental efficiency curves are shown in figure (8.5). They correspond to gratings nominally having blaze angles of 8.63^{o}, and line densities of 300, 600 and 1200/mm. The curves for the grating having 1200 lines/mm differ considerably from the curves for the other two gratings. For S polarization, the angular discrepancy reaches 3.5^{o}. Such substantial deviations suggest that the profile of the 1200 line/mm grating differs considerably from those of the other two gratings.

The theoretical curves of figure (8.6) correspond to a grating having a blaze angle of 8.63^{o}, and a groove apex angle of 120^{o}. For P polarization, good quantitative agreement exists between the theory and measurements made on the gratings having 300 and 600 lines/mm. The peak theoretical and experimental efficiencies differ by only 2%. For S polarization, all the theoretical anomalies correspond to features of similar form on one or the other of the experimental curves. It is again noticeable that the measured anomalies are always wider and are generally weaker than would be expected from the predictions of the infinite-conductivity theory. The computed peak efficiency associated

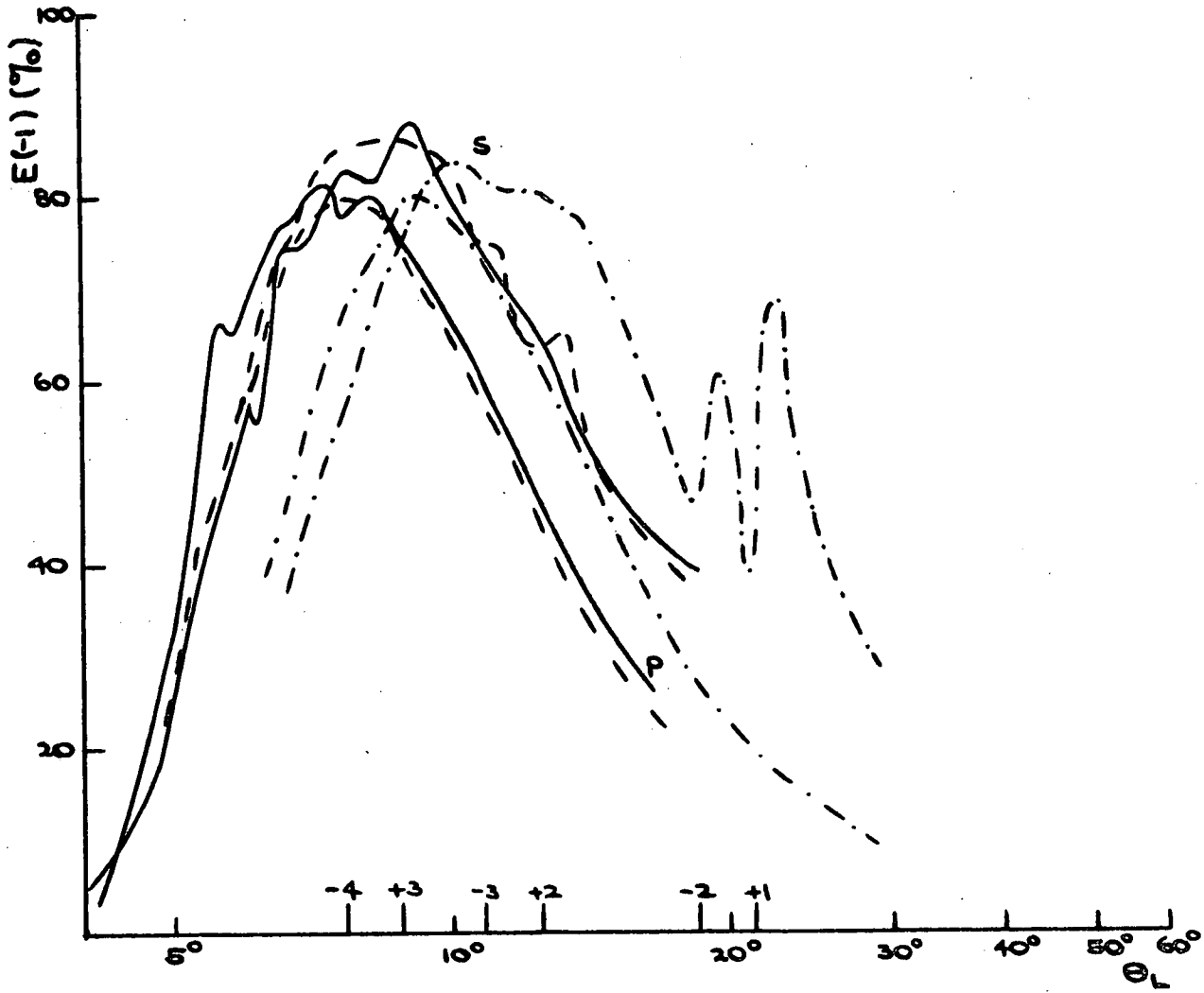


Figure (8.5). These curves show the variation of measured relative efficiency values in P and S polarized light with the Littrow diffraction angle, for three gratings each having a nominal blaze angle of 8.63^{0}, used with an angular deviation of -8.0^{0} in order -1. The solid lines correspond to the grating having 300 grooves/mm, while the dashed lines refer to the grating having 600 grooves/mm. The third grating has 1200 grooves/mm. (After E.G. Loewen.)

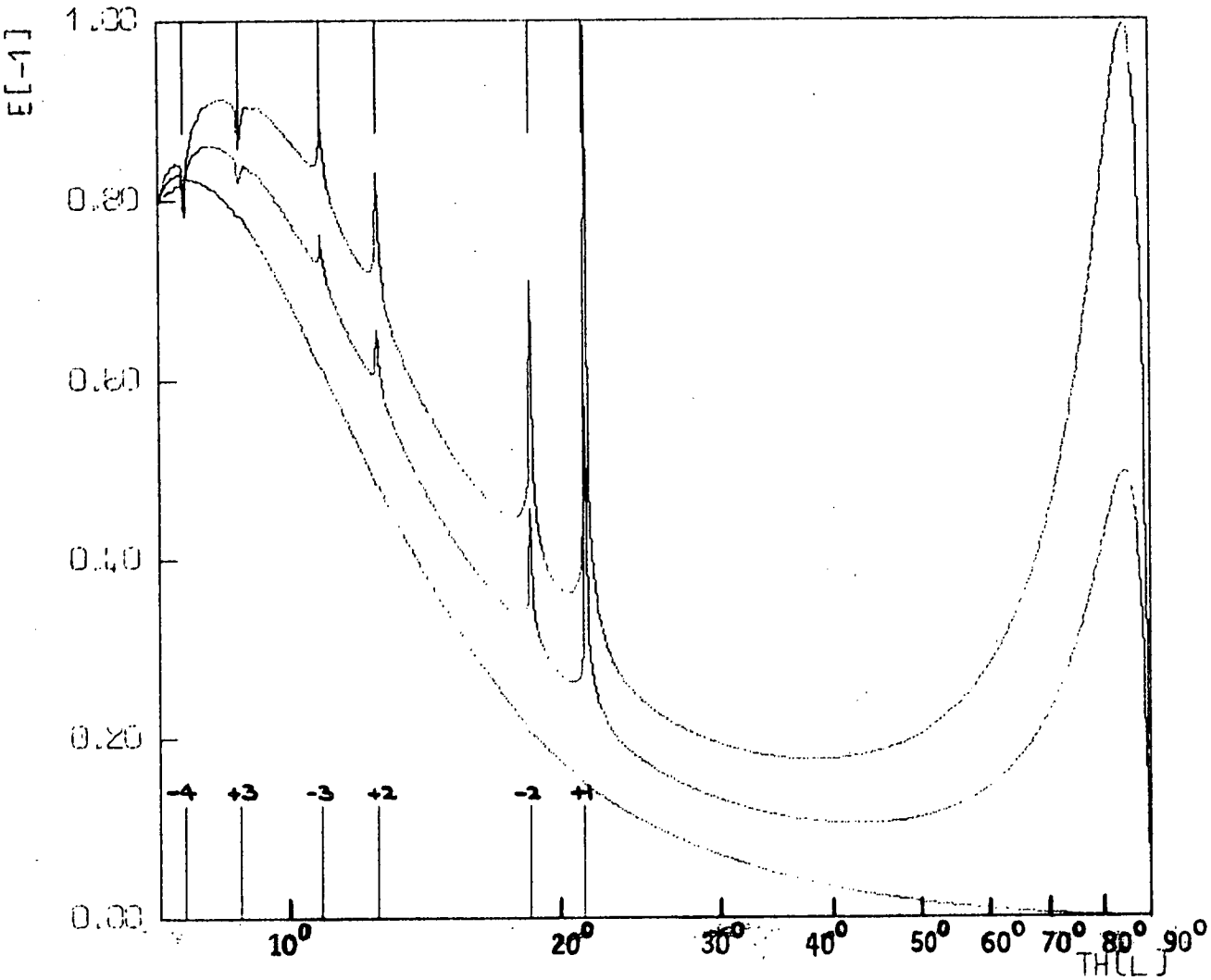


Figure (8.6). Theoretical efficiency curves are given for P, S and U polarized light, for a triangular profile grating having a blaze angle of 8.63° and a groove apex angle of 120°, used with an angular deviation of -8.0° in order -1. The abscissa is the Littrow diffraction angle, which is plotted on a logarithmic scale.

with the S polarization blaze differs by only 3% from the largest measured value.

In summary, for P polarization quantitative agreement exists in general between theoretical and experimental efficiency values. For S polarization, the theory successfully predicts the general shape of the experimental efficiency curve, and also indicates the correct form for the observed anomalies. However, substantial quantitative differences occur between theory and experiment, and also between nominally equivalent experimental efficiency curves. Evidently, for these three triangular profiles the form of the efficiency curve is more sensitive to profile variation for S polarization than for P polarization.

8.3. A TRIANGULAR PROFILE WITH LAND

Although rulers of gratings generally seek to impress a triangular groove shape on the blank, their efforts sometimes result in the formation of gratings having a residual amount of unruled surface. (A groove profile containing an unruled element, i.e. land, is shown in figure (1.7).) It is therefore desirable to know the effects of this profile defect on the form of the grating's efficiency curve.

R. Deleuil (8.2) has made measurements at millimetre wavelengths which show the effects of the introduction of land on the P polarization spectral performance of a triangular profile grating. His results are reproduced in figure (8.7). Corresponding theoretical curves obtained using the formulism of Pavageau and Bousquet are shown in figure (8.8).

The good qualitative agreement existing between theory and experiment is quite evident. Both figures show that introducing land

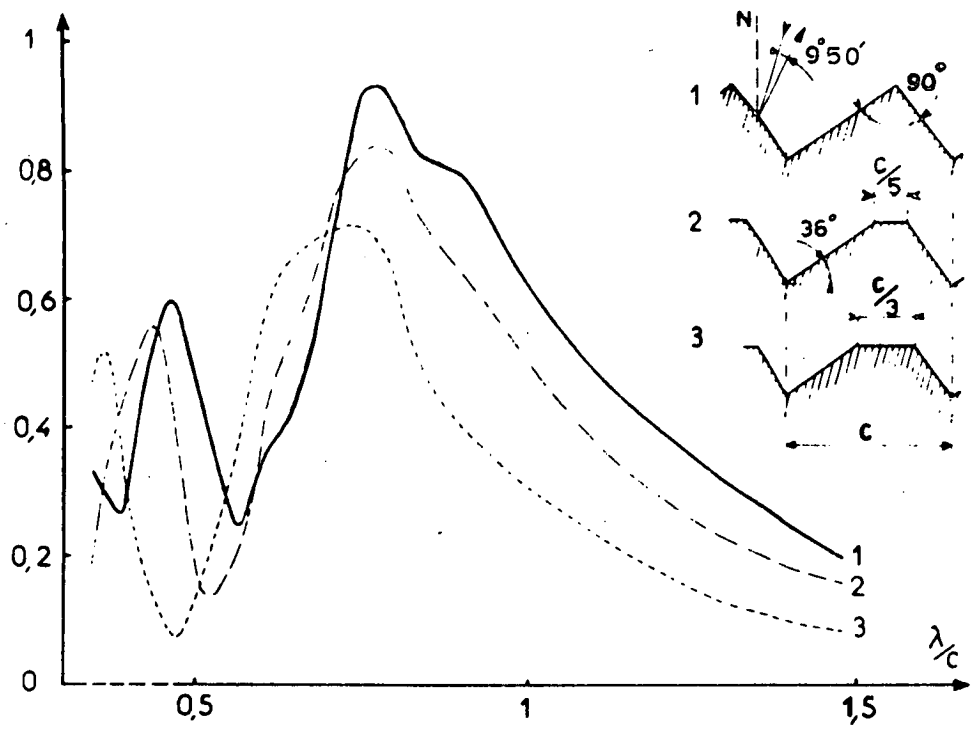


Figure (8.7). These experimentally-determined P polarization efficiency curves for the order +1 show the effects of the introduction of land into the profile of an echelette grating, which is used with a constant angular deviation of $9^{o}50'$ in the order +1. (After R. Deleuil.)

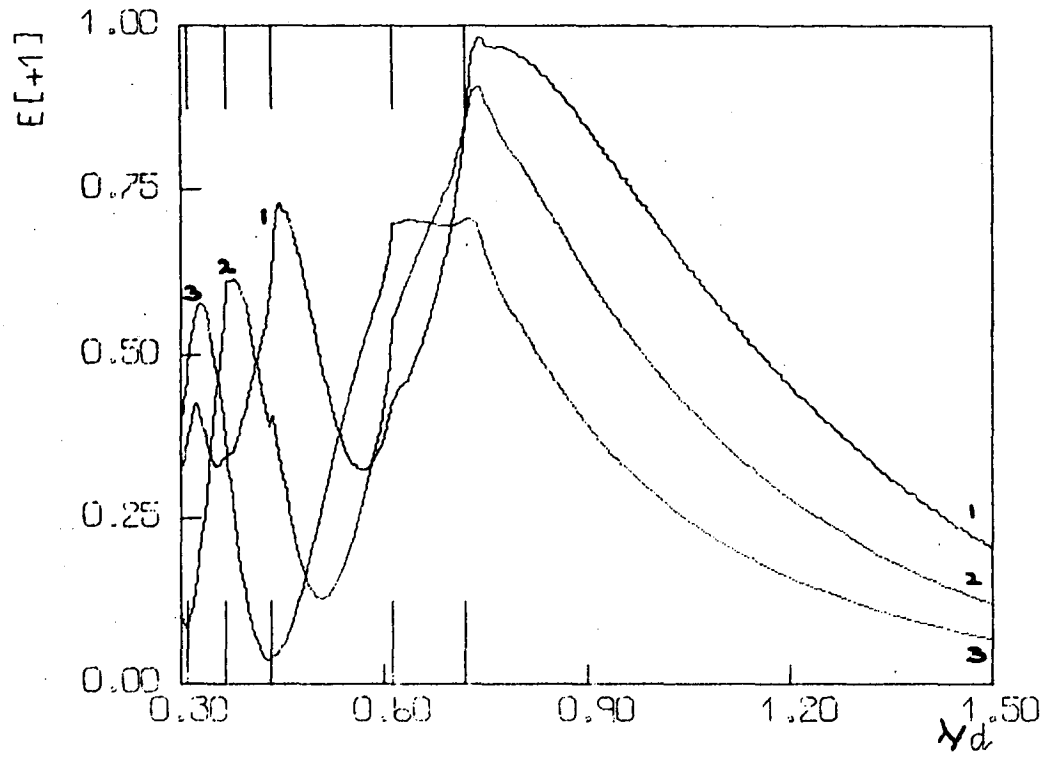


Figure (8.8). These efficiency curves are theoretical counterparts to the experimental curves of figure (8.7).

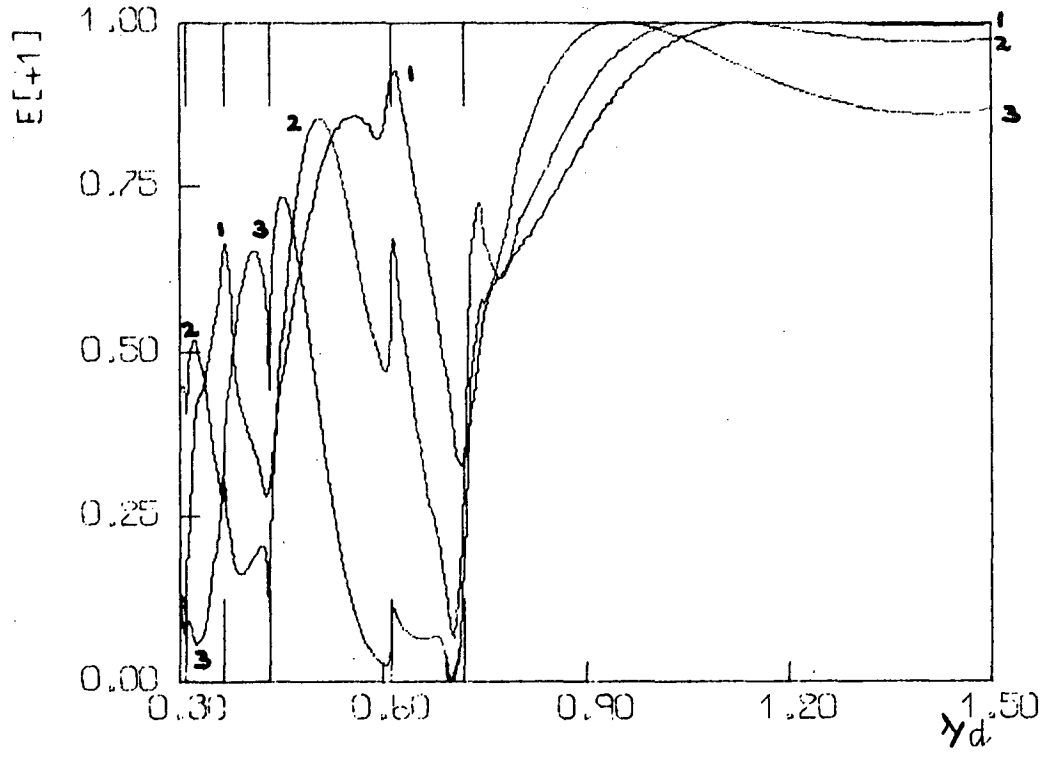


Figure (8.9). These theoretical efficiency curves correspond to the same data as those of figure (8.8), except that S polarization behaviour is shown.

not only results in a general lowering of the first order diffraction efficiencies, but also causes a bodily shift of the efficiency curve towards shorter wavelengths. (Theoretical efficiency curves given by R. Petit (8.3) also demonstrate these effects, although to a less marked degree.)

For the triangular profile 1, the theoretical and experimental curves are in qualitative agreement. The experimental curve has a more pronounced levelling-out of its gradient just after the blaze peak than has the theoretical curve. Both the wavelength locations of the main and subsidiary peaks are accurately predicted by the theory. The largest efficiency difference (12%) between theory and experiment occurs at the subsidiary peak.

The two curves for the profile 2(which has land occupying one fifth of each period) differ by only a few per cent on the long wavelength side of the blaze. However, the theoretical curve rises 6% higher to reach its blaze peak than does the experimental curve. The wavelength at which the theoretical efficiencies peak is slightly lower than the corresponding value for the experimental results. Similar deviations between the two curves occur at the subsidiary peak.

For the third profile, theory and experiment differ by only a few per cent everywhere through the blaze region and on its long wavelength side. The theory successfully predicts the formation of a rather wider and lower blaze peak than was associated with the other two profiles. From figure (8.8), it can be seen that the blaze peak is in fact contained between the Rayleigh wavelengths for orders +2 and -1. The subsidiary peak of figure (8.8) is about 5% stronger than that of figure (8.7), and the two occur at slightly different normalized wavelengths (0.33 and

0.36, respectively).

Figure (8.9) gives theoretical efficiency curves corresponding to those of figure (8.8), except that the radiation incident on the gratings is now S polarized. The introduction of land into the grating profile does not cause a decrease away from 100% of the S polarization blaze efficiency. However, the drop away from the blaze peak on its long wavelength side increases with land-length. The blaze wavelength moves towards the Rayleigh wavelength for order -1 as the groove profile becomes more shallow. In the region where four orders propagate, the increase of land also causes a displacement of features of the efficiency curves towards shorter wavelengths.

In summary, for P polarization the theory predicts accurately the changes in form of the efficiency curve caused by the introduction of an unruled segment into the triangular profile. Theory and experiment are generally in quantitative agreement in the wavelength region in which only two orders propagate. The theory gives main and subsidiary blaze peaks which are too strong and are located at wavelengths which are slightly too low. For S polarization, the theoretical results indicate that the strength of the blaze peak is much less affected by the land-length than is the case with radiation of the orthogonal polarization.

8.4. PROFILE DISTORTION AND DIFFRACTION ANOMALIES.

As mentioned in Chapter 3, the behaviour of gratings having a sinusoidal profile is now of great practical interest, since such gratings can be generated by holographic means. We have already given theoretical results displaying properties of diffraction anomalies of such gratings. The first experimental investigations of sinusoidal

profile anomalies have only recently been completed by M.C. Hutley (8.4, 8.5).

Figure (8.10) is taken from reference (8.5), and shows diffraction anomalies given by gratings having 730 lines/mm in S polarized light. From electronmicroscopic data, we may conclude that the grating profiles were close to sinusoids in form (8.5). Thus, these experimental curves correspond to those of figure (3.1)(a,c,e,f,g), in showing the effect of increasing groove amplitude on the double anomaly in order +1, associated with the grazing orders -3 and +2.

Theory and experiment agree in so far as the shape of the anomalies is concerned. The minimum formed for shallow grooves widens with increasing amplitude, and a local efficiency maximum appears at the Rayleigh wavelength. However, the calculated anomalies are very much narrower than the observed features. Hutley (8.5) compares the predicted and measured separations between the Rayleigh wavelength and the centre of the resonance minimum, for a grating of normalized amplitude 0.05, and shows that the two are approximately in the ratio 1 : 130.

It is clearly of importance that the cause or causes of such a large discrepancy in the wavelength scale of anomalies be isolated. Hutley (8.5) suggests three possible causes. The first is the finite conductivity of the silver surface coating of the experimental gratings. We have previously mentioned in this chapter that we would expect a broadening of resonance features because of the influence of finite conductivity. However, a theoretical investigation of conductivity effects cannot be undertaken with the programs presently available to the author. The second possible cause lies in deviations of the profiles of the experimental gratings from the sinusoidal form. Several experimental investigations (8.6, 8.7, 8.8, 8.9) have indicated

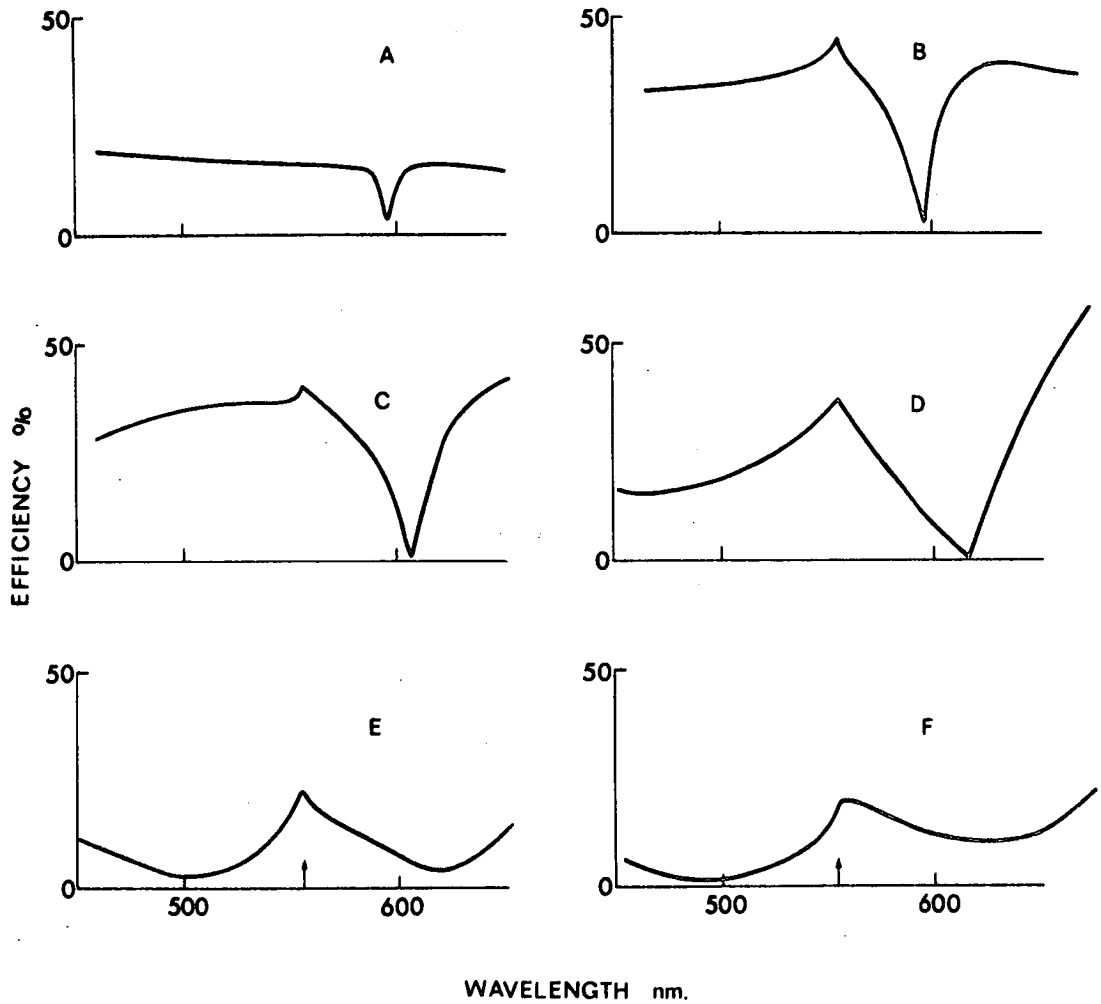


Figure (8.10). The variation with groove depth of the anomalies of silver coated gratings having 730 lines/mm is shown. The efficiency curves refer to the order +1, which was used with a constant angular deviation of 48.3^{o}. The normalized amplitude values of the experimental gratings are respectively:

A - 0.025, B - 0.05, C - 0.06, D - 0.08, E - 0.11, F - 0.12.

(After M.C. Hutley.)

that the form of Wood anomalies may be quite sensitive to profile changes. However, there exists no theoretical confirmation of this. As a third possibility, Hutley suggests that, although the gratings were used in a Littrow configuration, their anomalies associated with the two grazing orders -3 and +2 may not in fact have merged. This would be in keeping with the experimental observations of Stewart and Gallaway (8.8), who reported a reluctance of anomalies to merge. If the two anomalies did remain as separate entities, this would widen the anomalous wavelength region. The assumptions made in order to perform computations of grating efficiency may have a more profound effect on the form of the theoretical curves when two anomalies approach close together and interact, rather than when each anomaly is formed in isolation from the other.

In this section, we will investigate the second and third suggested causes. It will be shown that profile distortion can have an important effect on the form of Wood anomalies, and that the agreement between theory and experiment is better when anomalies are well separated rather than in close proximity to each other.

8.4.1. Distortion and the (-3,+2) Anomaly.

Figure (8.11) is taken from reference (8.5). It shows S polarization efficiency curves measured for a grating having 730 lines/mm and normalized groove amplitude 0.08, used in a Littrow mounting in the order -1. After efficiency measurements had been made on the aluminium-coated grating, a layer of silver was deposited on its surface and the measurements were repeated. The width of the resonance peak is larger for the silver coating, despite the fact that silver has a higher conductivity than has aluminium. (J. Strong (8.10) has given efficiency

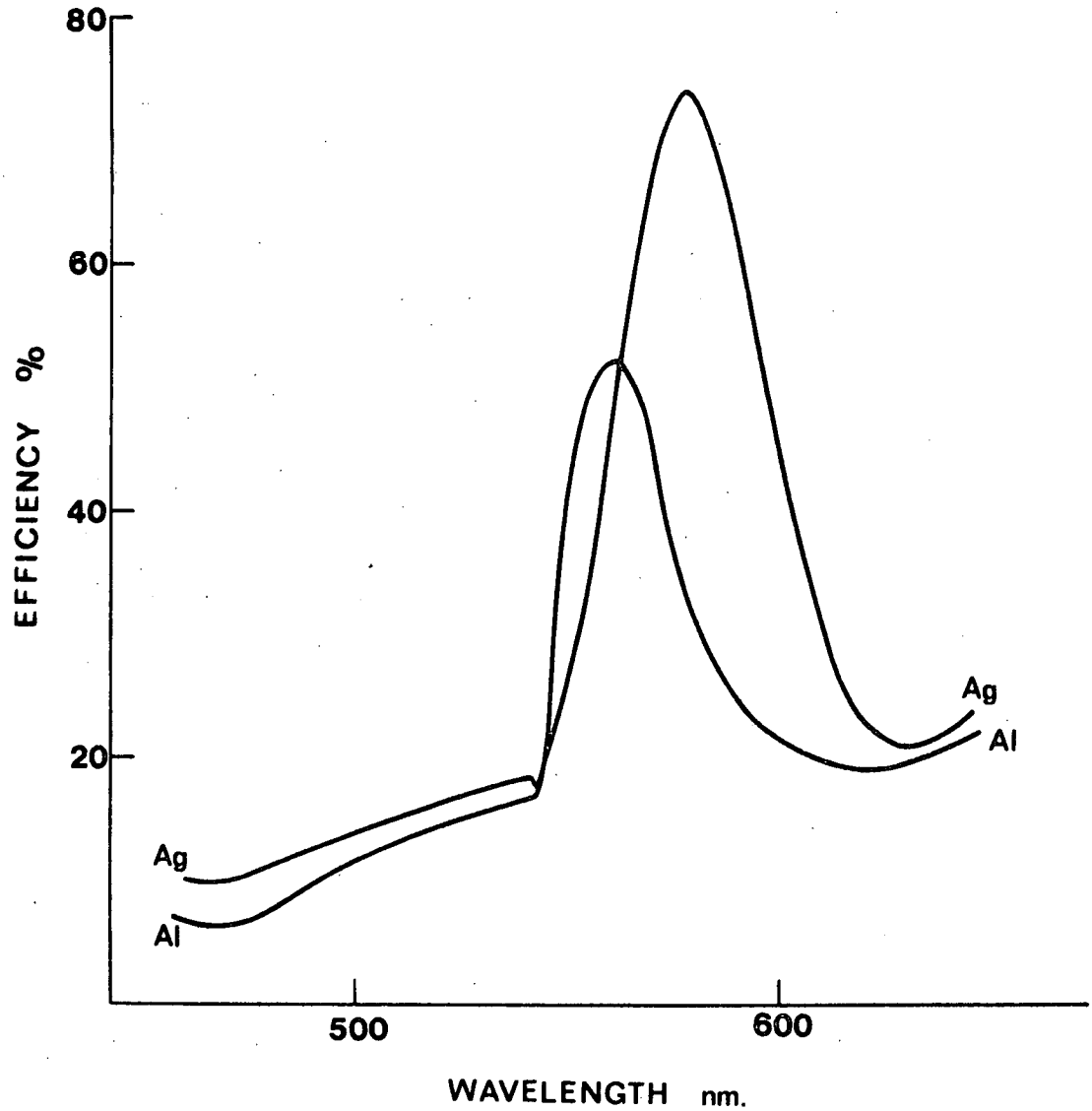


Figure (8.11). S polarization efficiency curves are given for a grating with normalized groove amplitude 0.08, covered successively with surface coatings of aluminium and silver. The curves refer to the order -1, which is used in a Littrow configuration. (After M.C. Hutley.)

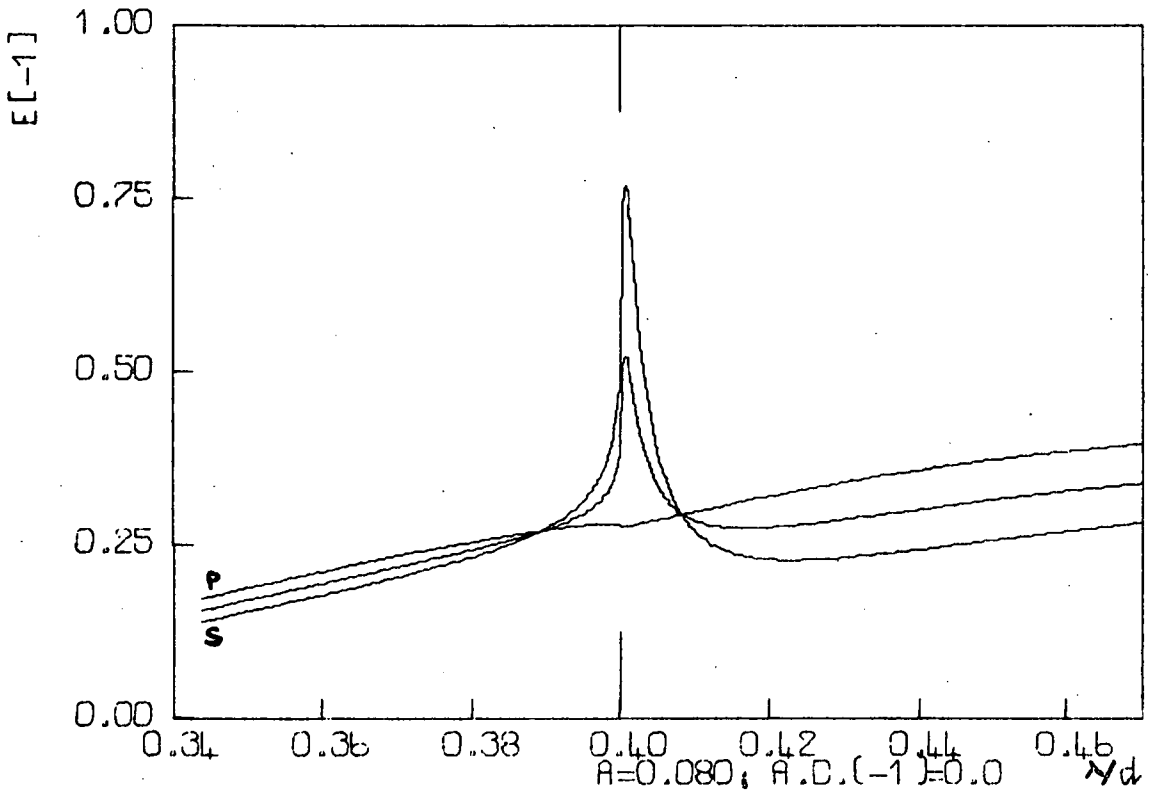
curves which confirm those of figure (8.11).)

Figure (8.12) shows theoretical efficiency curves for orders -1 and +1, for a grating having a sinusoidal profile with A/d = 0.08. The curve (a) for order -1 and S polarization has a resonance maximum of very similar strength to that given by the silver-coated grating. However, the half-width of the theoretical peak is much smaller than the corresponding values for the two experimental peaks. (A discrepancy of the order of five or six in the wavelength scale is involved.) Comparison of the theoretical curve (b) with the experimental curve of figure (8.10)D shows that the discrepancy of wavelength scale is much larger for the order +1 than for the order -1.

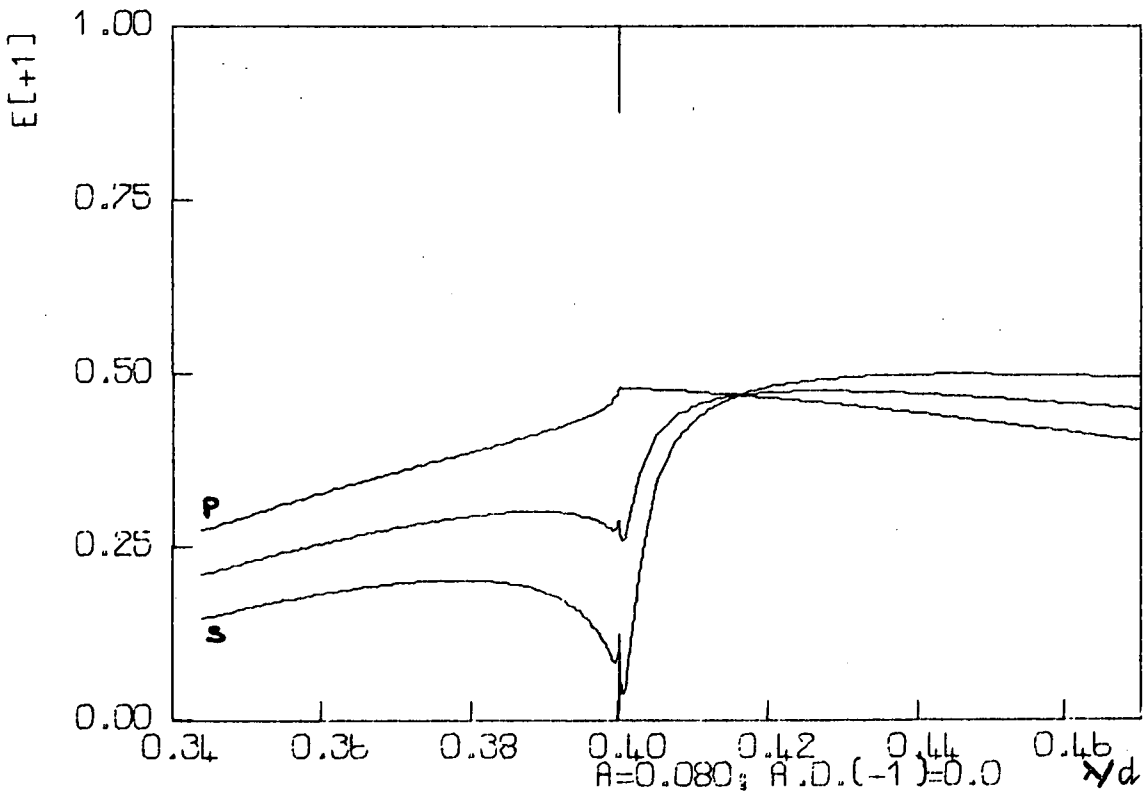
The effects of profile distortion on the S polarization anomalies in the orders -1 and +1 will now be investigated. The type of profile distortion considered will be that likely to arise because of the non-linearity of the photoresist (Shipley AZ1350) used in the fabrication of the experimental gratings.

Figure (8.13) shows five profile curves. Curve (1) is sinusoidal. Curves (2-5) incorporate increasing amounts of distortion. They have been obtained by functional multiplication of sinusoidal exposure patterns having various amplitude and pre-exposure factors with a sensitivity curve for Shipley AZ1350 photoresist measured by Beesley and Castledine (8.11). Non-inverted profiles have flatter groove tops than bottoms.

With increasing distortion from the sinusoidal profile in the non-inverted sense, the maximum efficiency associated with the resonance peak in order -1 decreases from 77% to 29% (for profile (5)). Furthermore, a steepening of the upper wavelength side of the peak causes



(a)



(b)

Figure (8.12). Theoretical efficiency curves for P, S and U polarizations are given for a grating with a sinusoidal profile having the normalized groove amplitude 0.08, used in a Littrow configuration in the order -1. (a) Order -1. (b) Order +1.

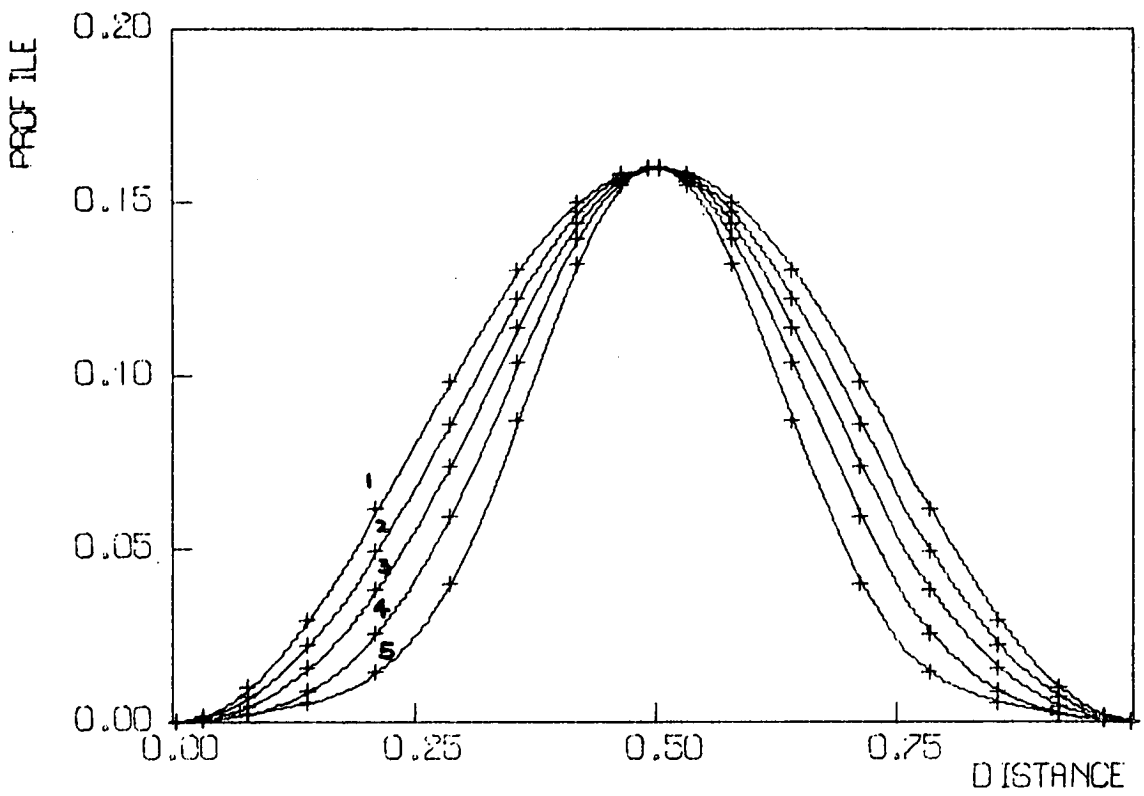
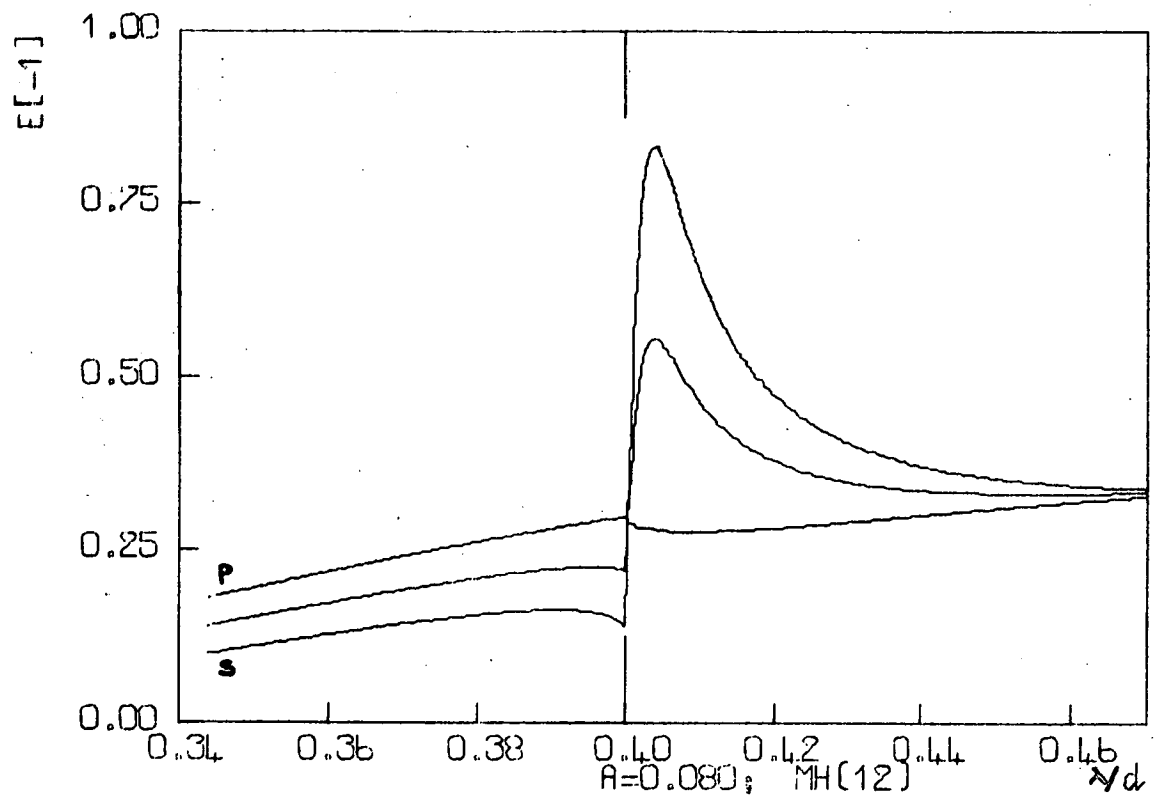


Figure (8.13). Five profiles having a common normalized groove depth (0.16) are shown. The profiles incorporate various amounts of distortion of the type given by the photoresist Shipley AZ1350. Profile (1) is sinusoidal. Non-inverted profiles have their top at ordinate zero, where the grooves are flattest.

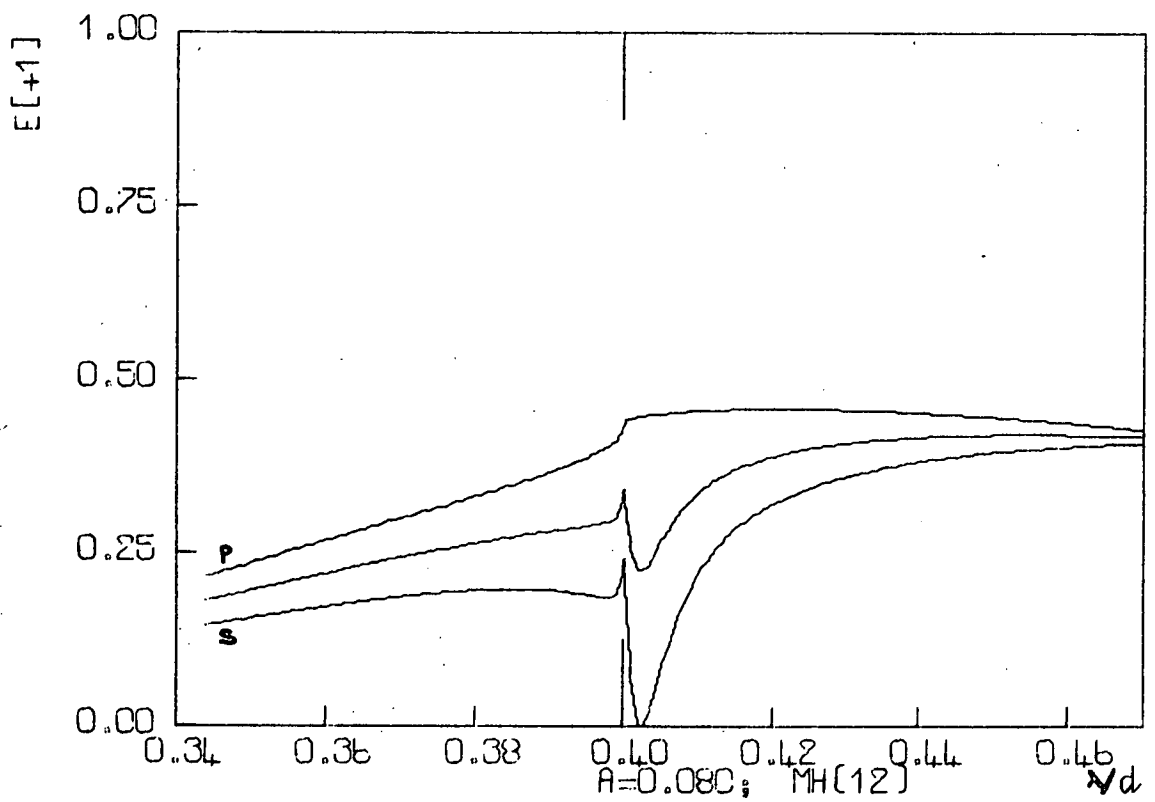
its half-width to decrease with increasing distortion. The agreement between the theoretical and experimental efficiency curves for order +1 is also worsened by the introduction of non-inverted profile distortion.

With increasing profile distortion in the inverted sense, the half-width of the resonance peak in order -1 increases, and the lower wavelength edge of the peak becomes less steep. A slight local efficiency minimum also forms at the foot of the peak, corresponding to the Rayleigh wavelength. Figure (8.14) gives efficiency curves for the orders -1 and +1, corresponding to the inverted profile (4). The theoretical S anomalies shown are in good qualitative agreement with experimental findings. Furthermore, the wavelength-scale discrepancy in order -1 has been reduced from five or six to a factor which is close to unity. In order +1, inverted profile distortion results in the resonance minimum deepening and moving away from the Rayleigh wavelength. The wavelength-scale discrepancy is reduced by distortion, but is still of the order of twenty for inverted profile (4).

Figure (8.15) shows experimental efficiency curves obtained by M.C. Hutley (8.5) in S polarized light, for a grating with a normalized profile amplitude of 0.14 and for silver and aluminium surface coatings. Theoretical efficiency curves in the order -1 for a sinusoidal grating having the same profile amplitude are given in figure (8.16). The agreement between theory and experiment is poor. The measured efficiency values at the lower end of the wavelength range are of the order of 40%. For sinusoidal gratings with normalized amplitudes in the range 0.12 to 0.16, the corresponding theoretical levels are of the order of a few per cent. Thus, variations of groove amplitude alone could not cause the discrepancies between theory and experiment.



(a)



(b)

Figure (8.14). These theoretical efficiency curves correspond to those of figure (8.12), except that the grating now has the inverted profile (4).

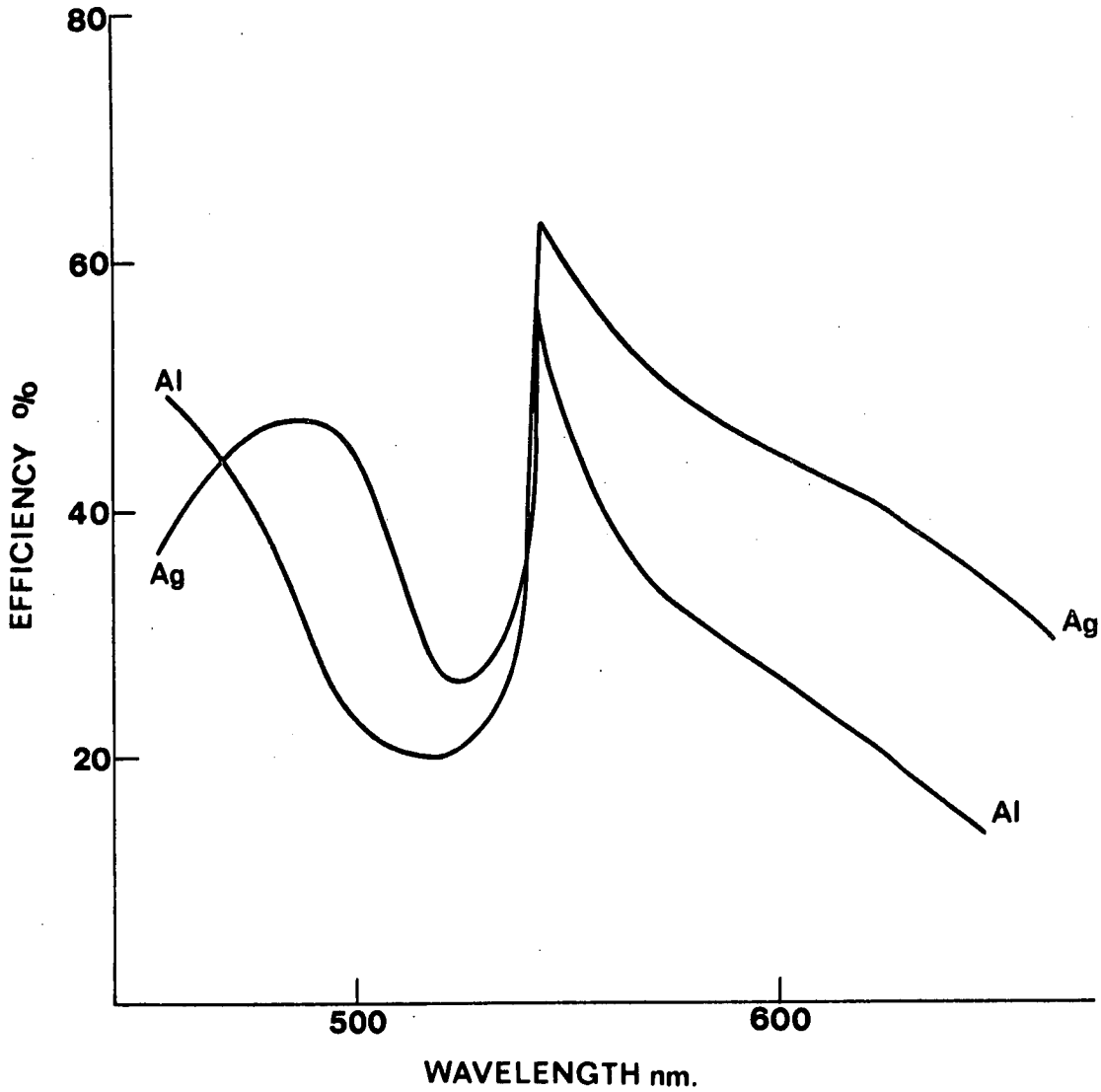


Figure (8.15). S polarization efficiency curves are given for a grating with normalized groove amplitude 0.14, covered successively with surface coatings of aluminium and silver. The curves refer to the order -1, which is used in a Littrow configuration. (After M.C. Hutley.)

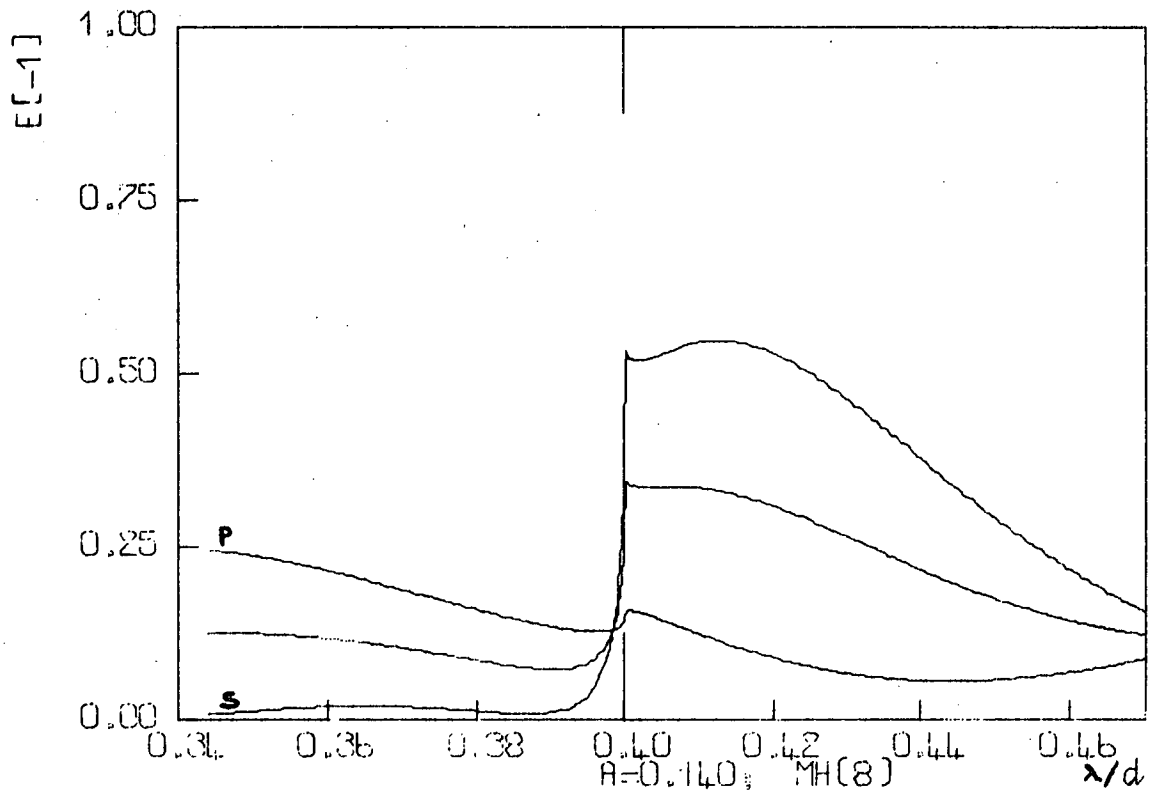


Figure (8.16). Theoretical efficiency curves for P, S and U polarizations are given for a grating with a sinusoidal profile having the normalized groove amplitude 0.14. The curves are for the order -1, which is used in a Littrow configuration.

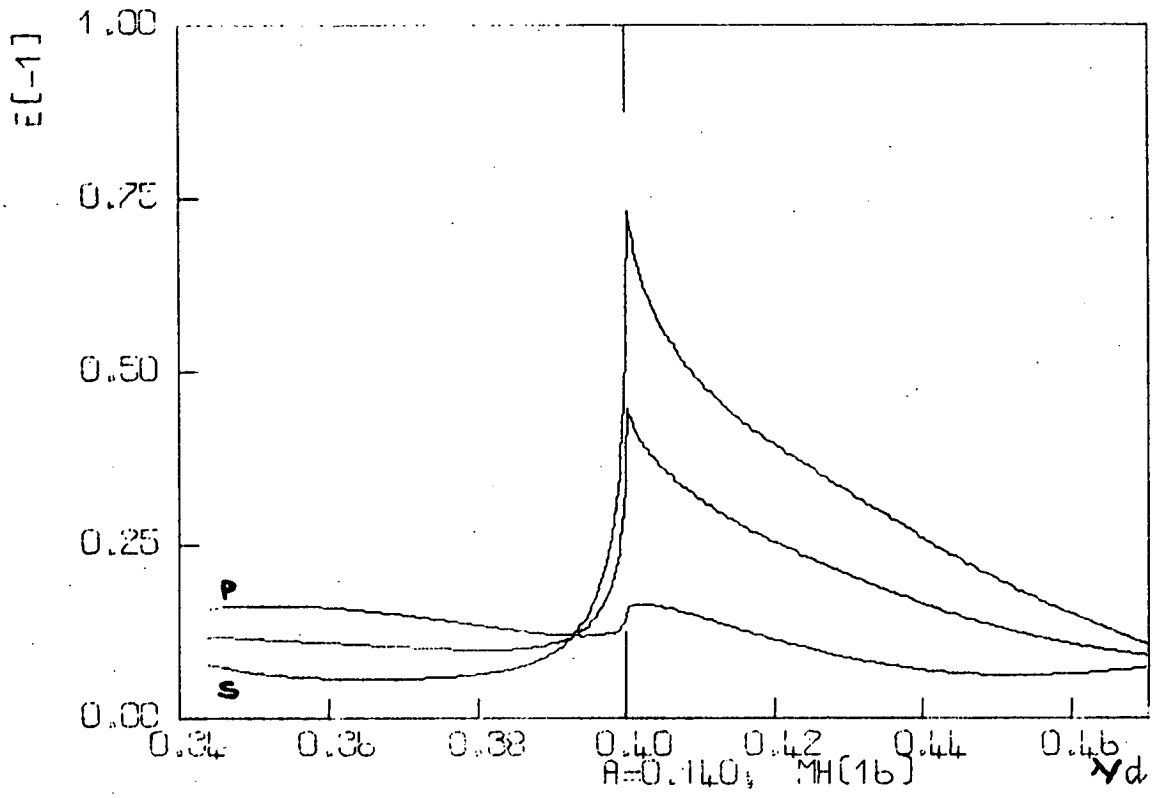
The theoretical efficiency curves of figure (8.17)(a) and (b) correspond respectively to the inverted profiles (2) and (3). The great profile sensitivity of this S polarization anomaly becomes quite evident when figures (8.16) and (8.17)(a) are compared. Profiles (1) and (2) are so similar that the difference between them could not be detected using current mechanical and electronmicroscopic observation techniques. However, the anomalies formed by the two profiles are quite different.

The S efficiencies in order -1 for wavelengths below λ_R increase steadily with inverted profile distortion. The strength of the efficiency jump near λ_R also increases, so that the anomaly takes on a pure branch-point nature, with the wide peak occurring after the Rayleigh wavelength for the sinusoidal profile ceasing to be evident. The inverted profile (3) has been chosen as giving the best agreement between theory and experiment for this anomaly. A comparison of figures (8.15) and (8.17)(b) indicates that there is good qualitative agreement between the measured and calculated anomalies. (The use of inverted profile distortion also ensures that theory and experiment are in qualitative agreement with regard to the S polarization anomaly formed in order +1 by the silver-coated grating.)

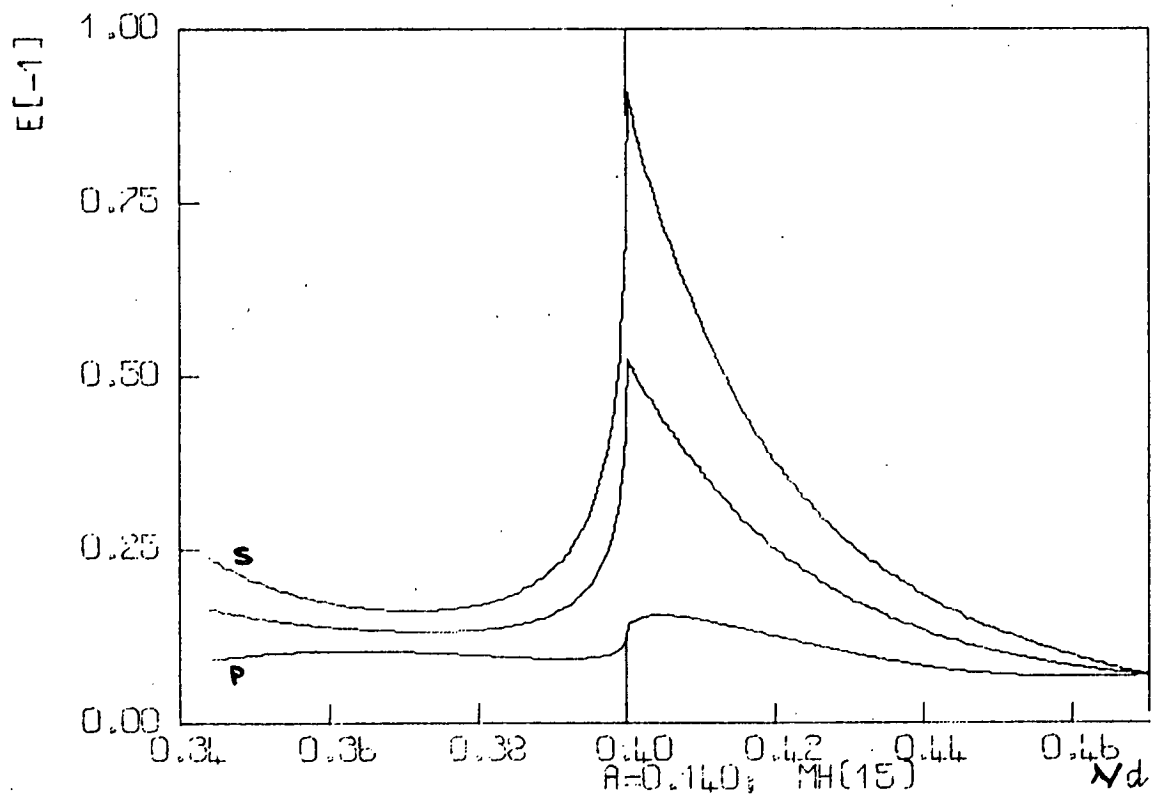
8.4.2. Amplitude Dependence of the (-2,+1) Anomaly.

In figure (8.18), taken from reference (8.5), experimental efficiency curves are given which show the variation of the (-2,+1) P and S polarization anomalies with groove depth. The measurements were made on aluminium coated gratings having 1200 lines/mm, used in a Littrow mount in order -1.

For each of the five groove depths, the effects of profile distortion on both the P and S polarization anomalies in order -1 were



(a)



(b)

Figure (8.17). These theoretical efficiency curves correspond to those of figure (8.16), except that now the grating has a distorted profile. (a) Inverted profile (2). (b) Inverted profile (3). (The profile curves 2 and 3 of figure (8.13) have been multiplied by a constant factor, to give normalized groove depths of 0.28.)

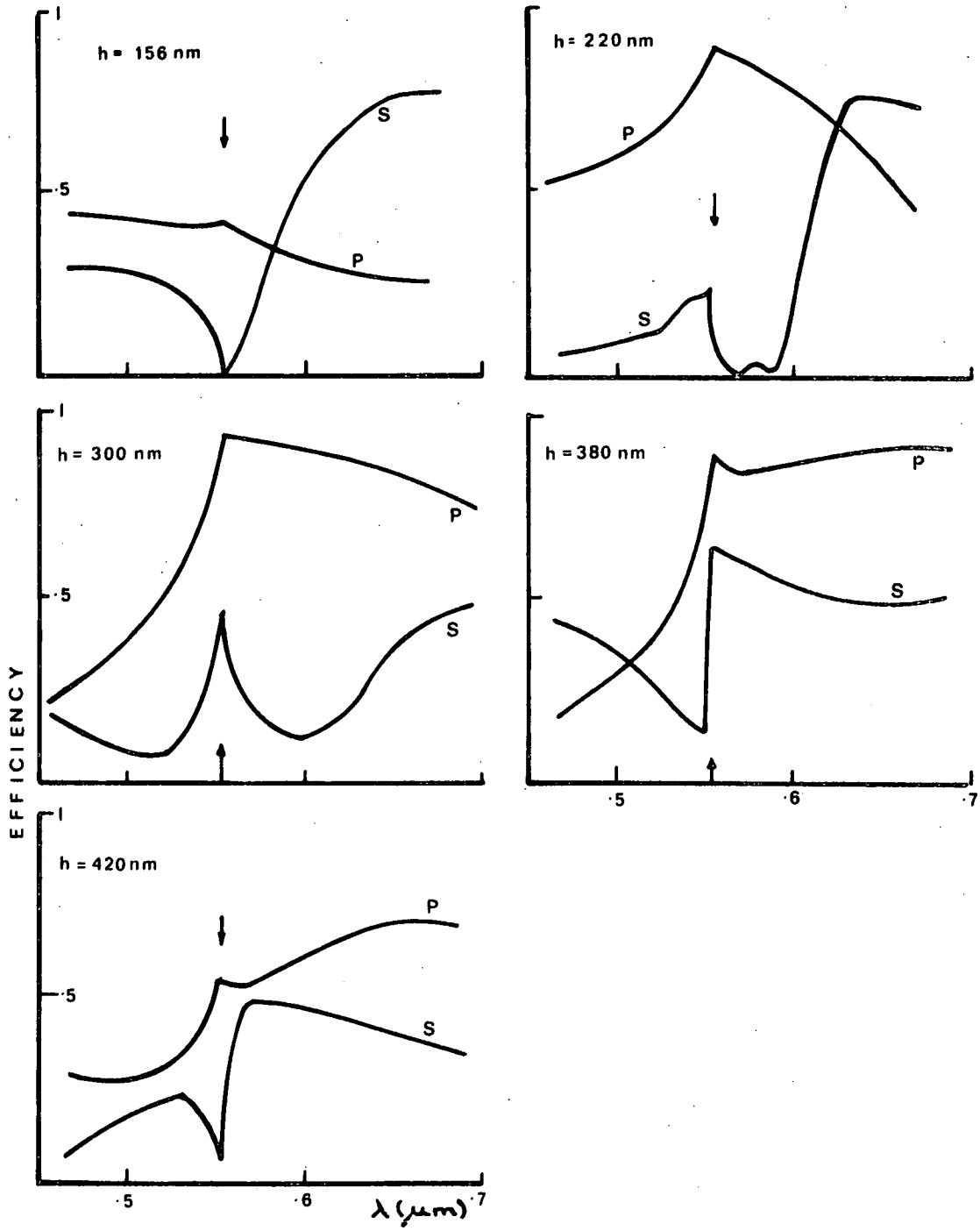


Figure (8.18). The variations with groove depth (h) of the (-2,+1) P and S polarization anomalies of gratings having 1200 lines/mm and an aluminium surface coating are shown. These experimental curves refer to the order -1, which is used in a Littrow mount. (After M.C. Hutley.)

considered, and one of the rescaled profiles (1-5) was chosen as giving the best agreement between theory and experiment.

Figure (8.19) shows theoretical efficiency curves for P, S and U polarizations, for a grating having the non-inverted profile (4) and a normalized profile depth of 0.1872. The P polarization curve is very similar in form to its counterpart in figure (8.18). The S polarization theoretical efficiency curve increases much too rapidly after passing through its minimum at the Rayleigh wavelength. The weak spike separating the efficiency jump from the rounded maximum is not seen on the experimental curve.

It was found to be quite difficult to obtain theoretical counterparts for the measured anomalies given by the grating having a groove depth of 220 nm (and a normalized groove depth of 0.264). Figure (8.20) shows theoretical efficiency curves for the non-inverted profile (3). For P polarization, quite good agreement exists between calculated and measured efficiencies. However, for S polarization there are major differences in form between the theoretical and measured anomalies. (It was verified that these differences could not be removed either by profile variation or by variation of groove depth within the range of its experimental error.)

For a normalized groove depth of 0.36, efficiency curves for three different profiles are given in figure (8.21). The non-inverted profile (3) gives efficiency values in best accord with measured values. For P polarization, theory and experiment are in quantitative agreement. For S polarization, while qualitative agreement exists, the theoretical peak formed near λ_R is too wide and strong.

For this groove depth, the transition from profile (1) to

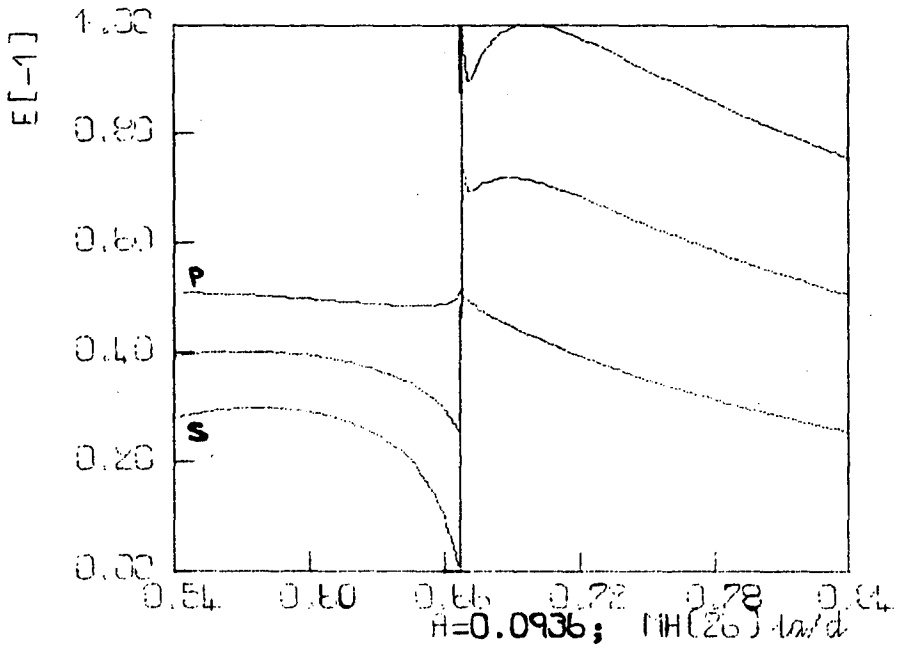


Figure (8.19). The theoretical efficiency curves correspond to a grating having a normalized groove depth of 0.1872 and the non-inverted profile (4), used in a Littrow mount in order -1.

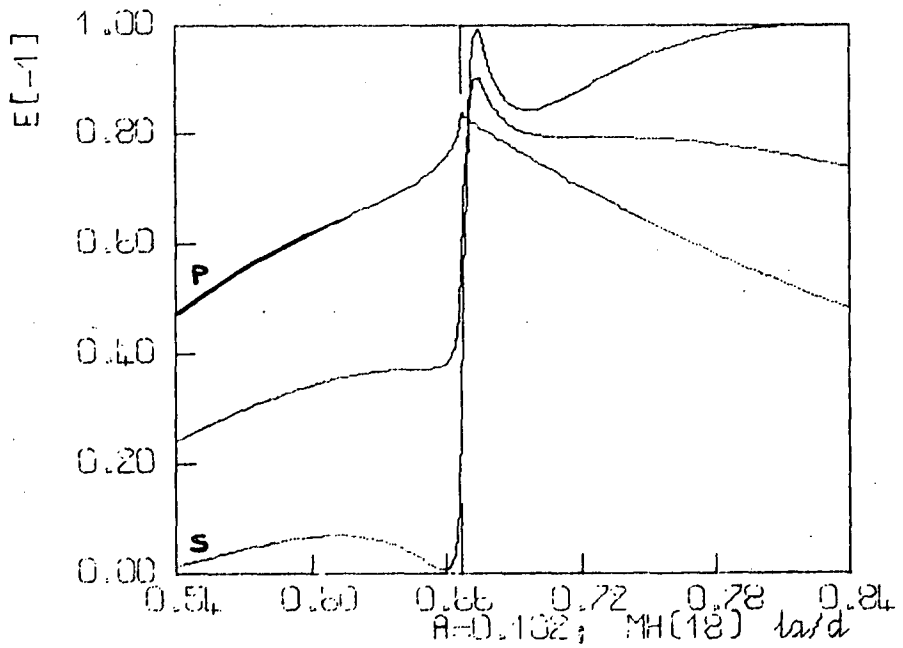
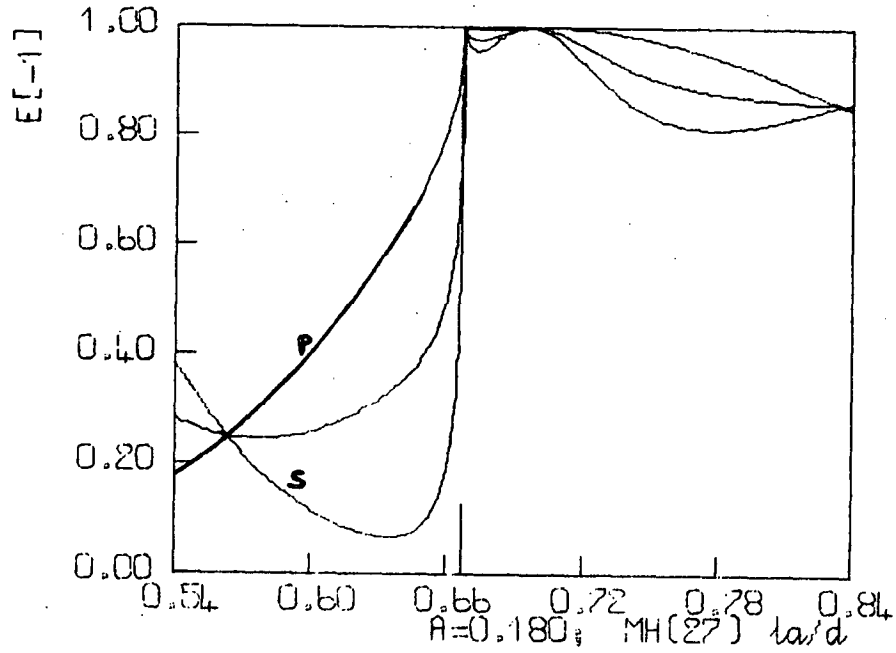
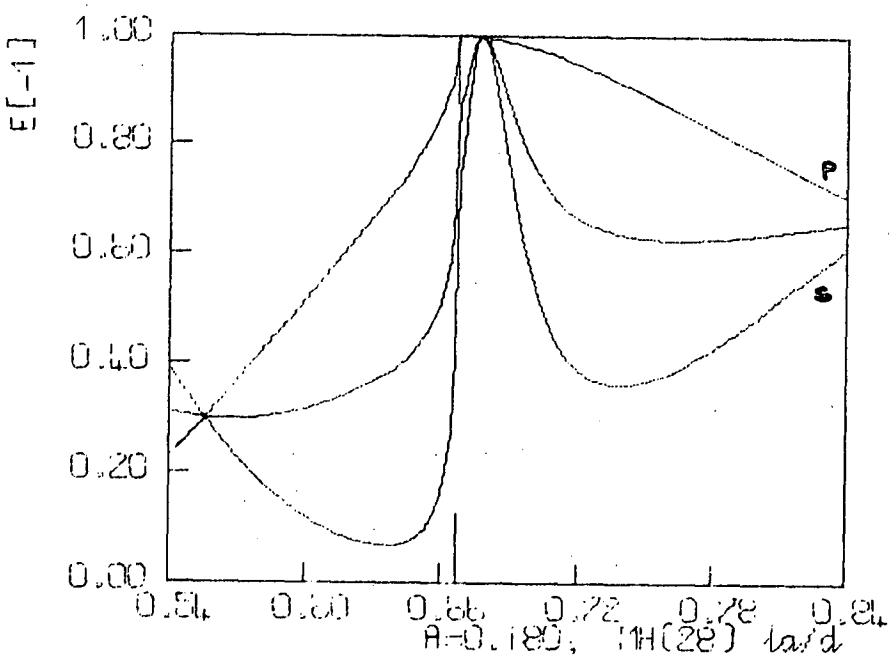


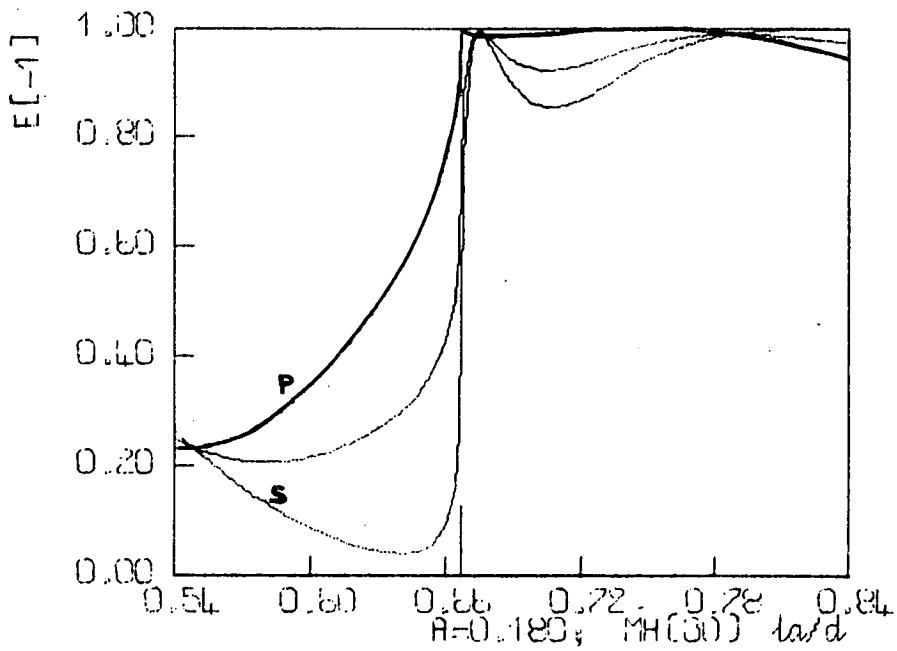
Figure (8.20). The theoretical efficiency curves correspond to a grating having a normalized groove depth of 0.264 and the non-inverted profile (3), used in a Littrow mount in order -1.


E[-1]
1.00
0.80
0.60
0.40
0.20
0.00
P
S
0.54
0.60
0.66
0.72
0.78
0.84
A=0.180; MH(27) la/d

(a)


E[-1]
1.00
0.80
0.60
0.40
0.20
0.00
P
S
0.54
0.60
0.66
0.72
0.78
0.84
A=0.180; MH(28) la/d

(b)



(c)

Figure (8.21). The three sets of theoretical efficiency curves correspond to gratings having a normalized groove depth of 0.36, used in a Littrow mount in order -1. (a) Sinusoidal profile. (b) Non-inverted profile (3). (c) Inverted profile (3).

non-inverted profile (3) causes a great change in the form of the S polarization efficiency curve. However, the curve is much less affected by the transition to inverted profile (3). For P polarization, the transition from profile (1) to non-inverted profile (3) causes a contraction of the wavelength extent of the blaze peak. The transition to inverted profile (3) widens the blaze region, just as if we had retained the sinusoidal profile shape but had deepened the groove depth.

Consider next the fourth graph of figure (8.18), that for a grating with groove depth 380 nm. The theoretical profile chosen as providing the best agreement with experiment is sinusoidal, with a normalized amplitude of 0.21. The P polarization curve of figure (8.22) is in good accord with the experimental data. The S polarization curve agrees reasonably well with experiment, although the efficiency values are initially too high, and the drop away after the steep upward edge at λ_R is too pronounced.

For a groove depth of 420 nm, inverted profile (2) was used to provide the theoretical counterparts shown in figure (8.23) for the measured curves. Once again, the P polarization calculations agree reasonably well with experiment. The Wood anomaly has moved well down the resonance peak, but is rather stronger on the theoretical curve than on the experimental curve. For S polarization, there is qualitative agreement between measurements and computations. The minimum at λ_R increases rapidly in strength with inverted profile distortion. The theoretical efficiency values on both sides of the minimum are too high.

8.4.3. Anomalies in Isolation.

Figure (8.24) is taken from reference (8.5). For each of the five gratings considered in the previous sub-section, the variation

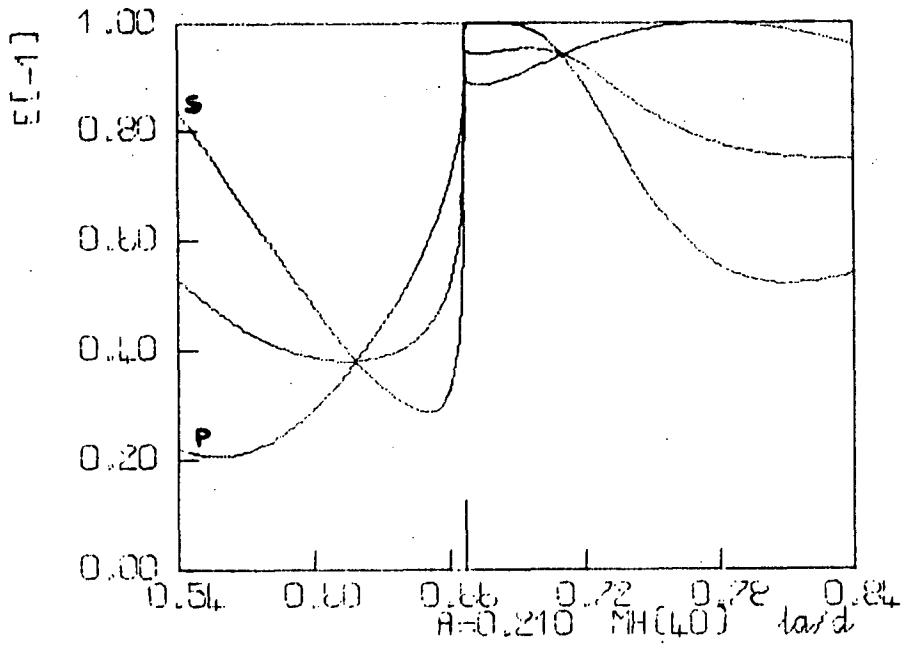


Figure (8.22). The theoretical efficiency curves correspond to a grating having a normalized groove depth of 0.42 and a sinusoidal profile, used in a Littrow mount in order -1.

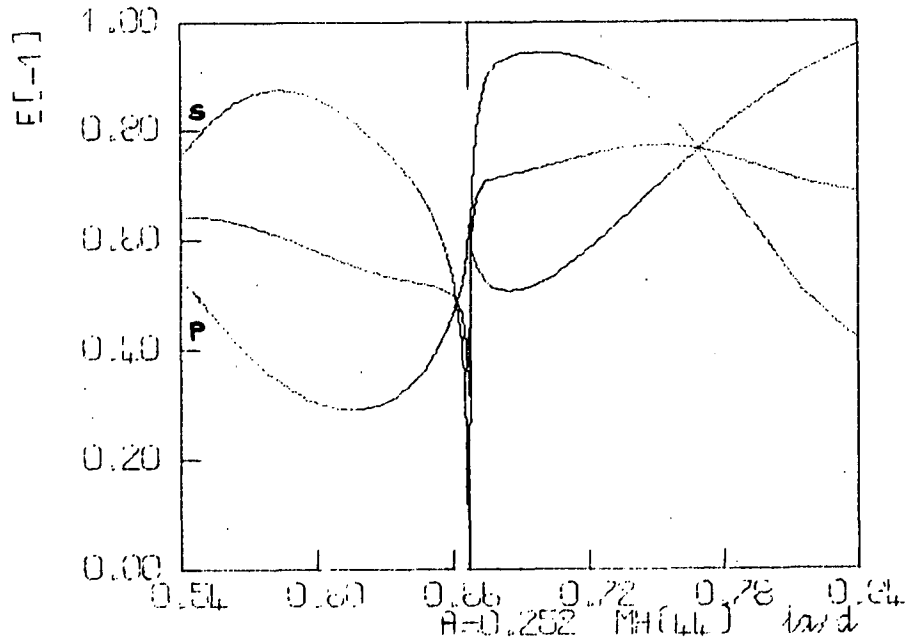


Figure (8.23). The theoretical efficiency curves correspond to a grating having a normalized groove depth of 0.504 and the inverted profile (2), used in a Littrow mount in order -1.

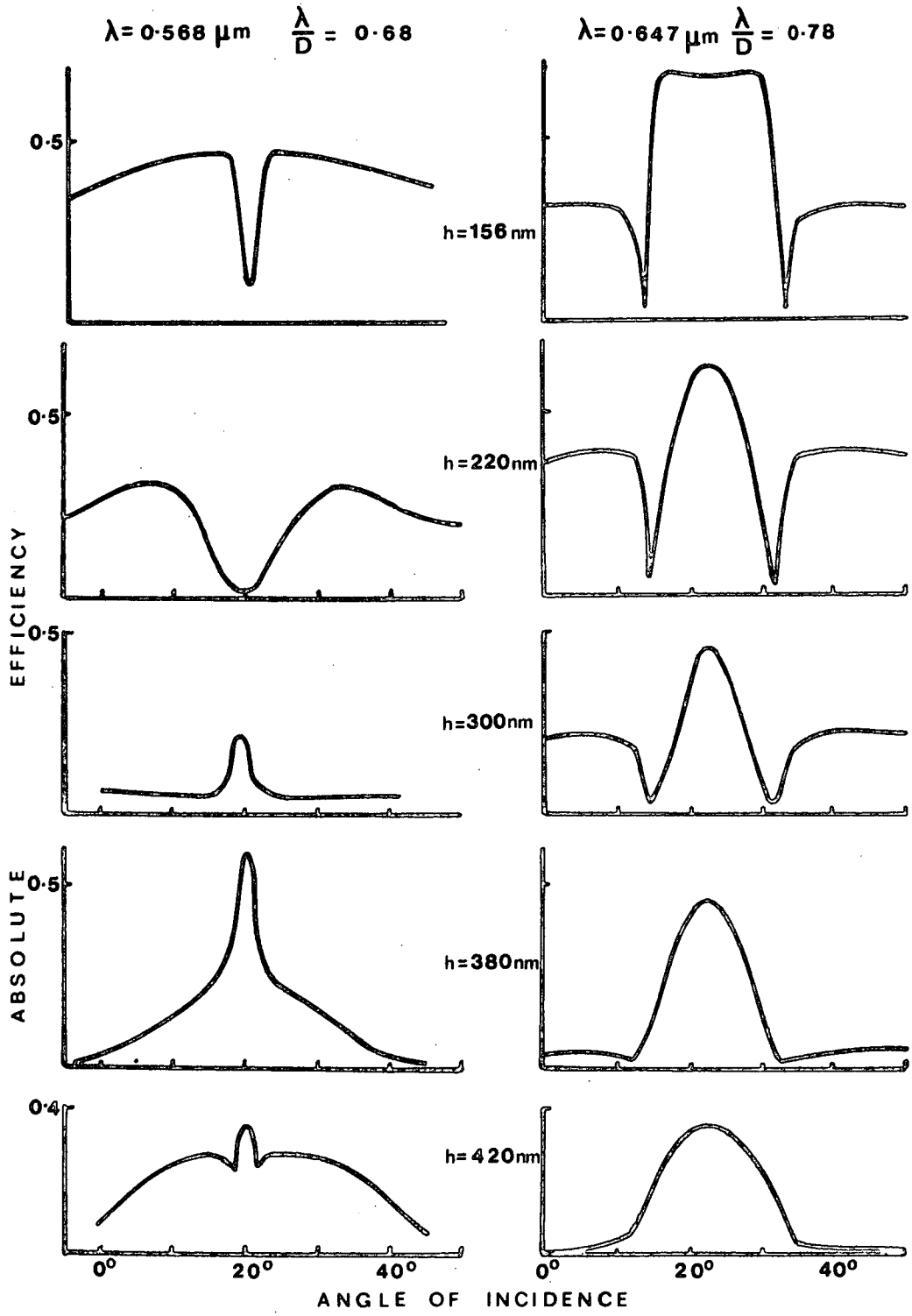


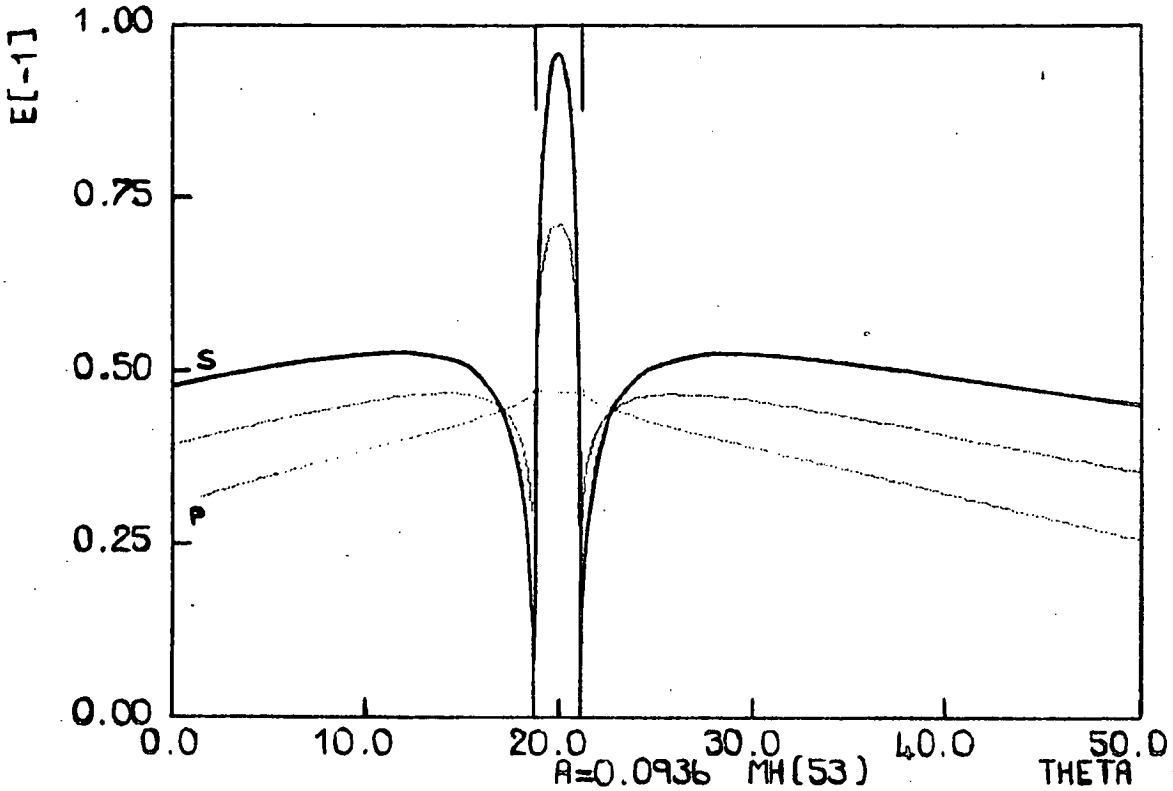
Figure (8.24). These experimental S polarization efficiency curves show the variation with groove depth (h) of the anomalies formed at two constant values of the normalized wavelength (0.68 and 0.78). The gratings have 1200 lines/mm and are aluminium coated. (After M.C. Hutley.)

of S polarization efficiency as a function of the angle of incidence is shown, for two fixed values of the normalized wavelength. For the lower of these, the two Rayleigh angles (18.7^{o} and 21.1^{o}) are quite close together, whereas for $\lambda/d = 0.78$ the Wood anomalies are well separated (the Rayleigh angles being 12.7^{o} and 34.1^{o}). All the efficiency curves show the approximate symmetry about the Littrow angle we would expect from the Reciprocity Theorem.

The five profiles chosen as giving the best agreement between calculations and measurements with regard to the form of the (-2,+1) anomaly were used to compute counterparts for the ten experimental curves. These are shown in figures (8.25-8.29).

Theoretical curves for the non-inverted profile (4) are given in figure (8.25). For $\lambda/d = 0.68$, the computed curve has a central maximum which was not observed experimentally. (This central maximum only disappears for the extreme distortion - profile (5) - in the inverted sense. The S polarization curve for inverted profile (5) has two sharp spikes corresponding to the Rayleigh angles, and an efficiency zero at the Littrow angle θ_L.) Despite this poor agreement for $\lambda/d = 0.68$, the qualitative accord between calculated and measured curves for $\lambda/d = 0.78$ is very good.

The theoretical curves of figure (8.26) correspond to non-inverted profile (3). For $\lambda/d = 0.68$, the computations again give a narrow maximum centred on θ_L, which is absent on the experimental curve. (This central maximum is also present on theoretical curves for profiles (1) and (3) inverted.) For $\lambda/d = 0.78$, the measured and computed curves are in qualitative agreement. The theoretical curve rises too steeply away from the minima, which causes the central maximum to be too wide



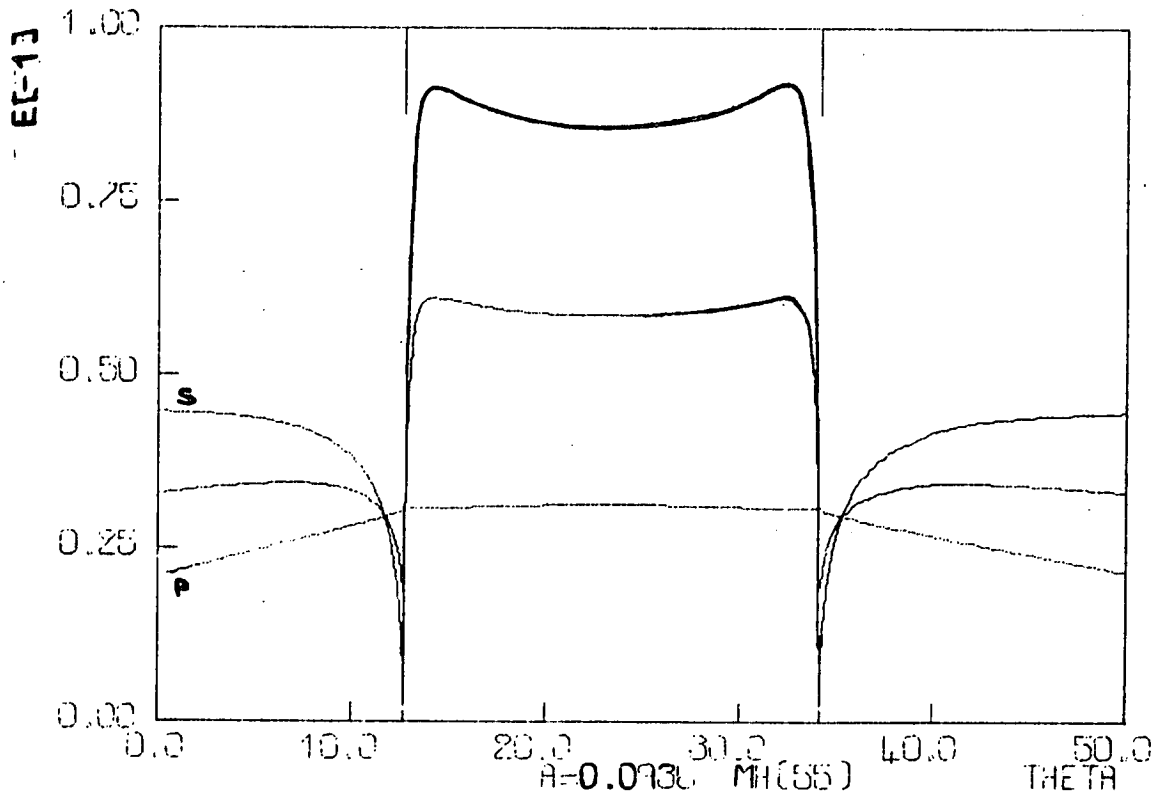


Figure (8.25). These theoretical efficiency curves refer to the behaviour of a grating having a normalized groove depth of 0.1872 and the non-inverted profile (4). The top and bottom graphs correspond respectively to the constant normalized wavelengths 0.68 and 0.78 (as is the case for the four following figures).

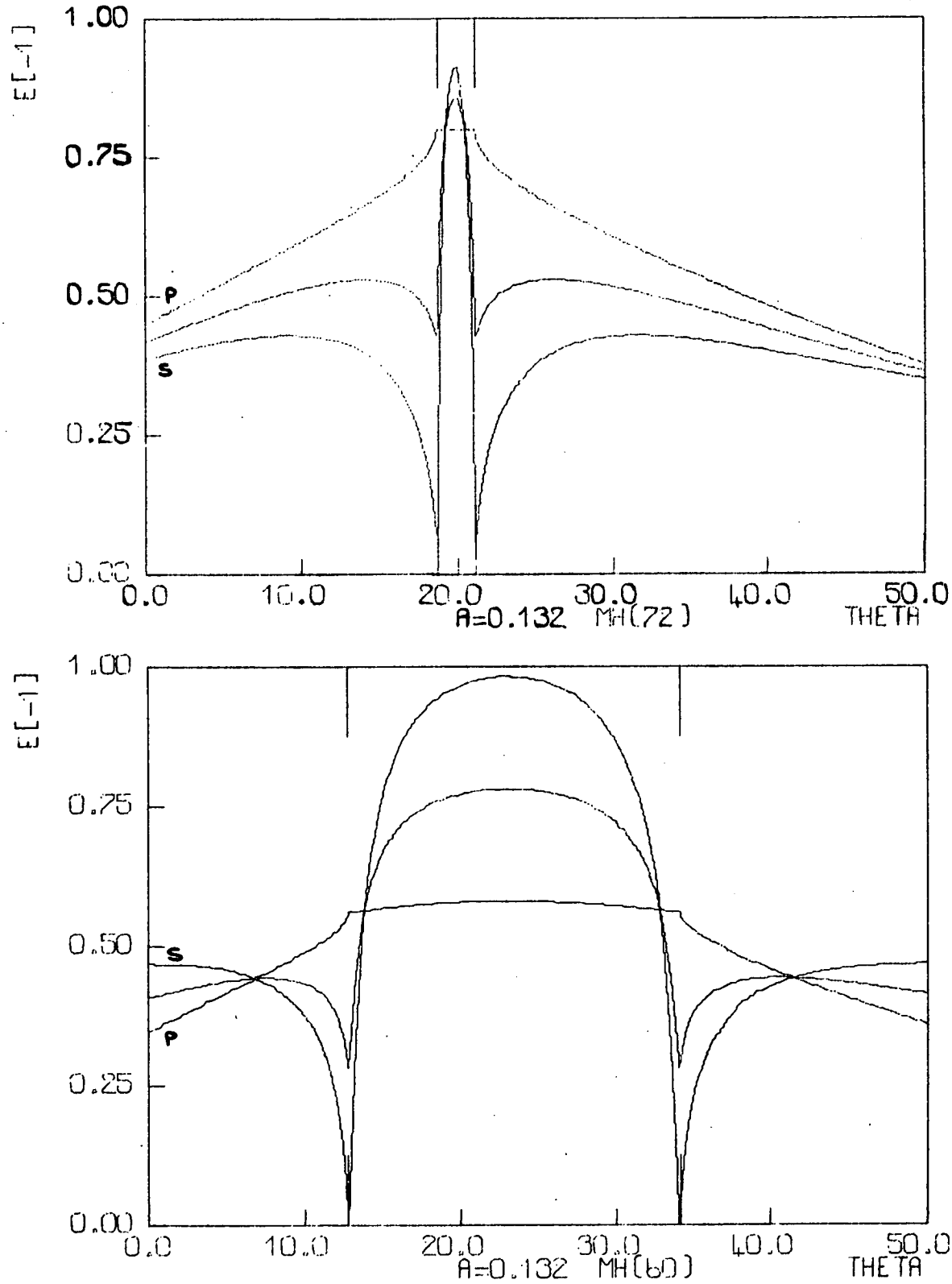


Figure (8.26). These theoretical efficiency curves refer to the behaviour of a grating having a normalized groove depth of 0.264 and the non-inverted profile (3).

as well as being too strong.

For the grating of groove depth 300 nm used with radiation of normalized wavelength 0.68, qualitative agreement exists between theory and experiment, despite the proximity of the two anomalies. The central maximum of the theoretical curve is much too strong. For $\lambda/d = 0.78$, qualitative accord also exists. The theoretical gradients around the Rayleigh angles are much steeper than the experimental gradients.

In the case of the next grating (that having a groove depth of 380 nm), there is good qualitative agreement between theory and experiment for both values (0.68 and 0.78) of the normalized wavelength. For the lower of these, the theoretical efficiency peak at θ_L is too strong. For the larger wavelength, the theoretical gradients near the Rayleigh angles are too great, which results in too wide a central maximum.

For the grating having the deepest grooves and with $\lambda/d = 0.68$, there is quite good qualitative accord between theory and experiment. The Rayleigh angles on one curve correspond to steep edges following almost horizontal sections; on the other, weak minima precede the edges. Both computations and measurements give edges which are much shorter than for the previous depth. For $\lambda/d = 0.78$, again the agreement is reasonable. As before, the Rayleigh angles on the theoretical curve correspond to much steeper gradients than were observed in the experimental situation.

For the data of figure (8.18), then, the agreement between theory and experiment is generally very good for P polarization. The agreement is less satisfactory for S polarization, particularly for shallower groove depths.

For the data of figure (8.24), the theory seems much less successful near the wavelength where the anomalies coincide ($\lambda/d = 0.66667$),

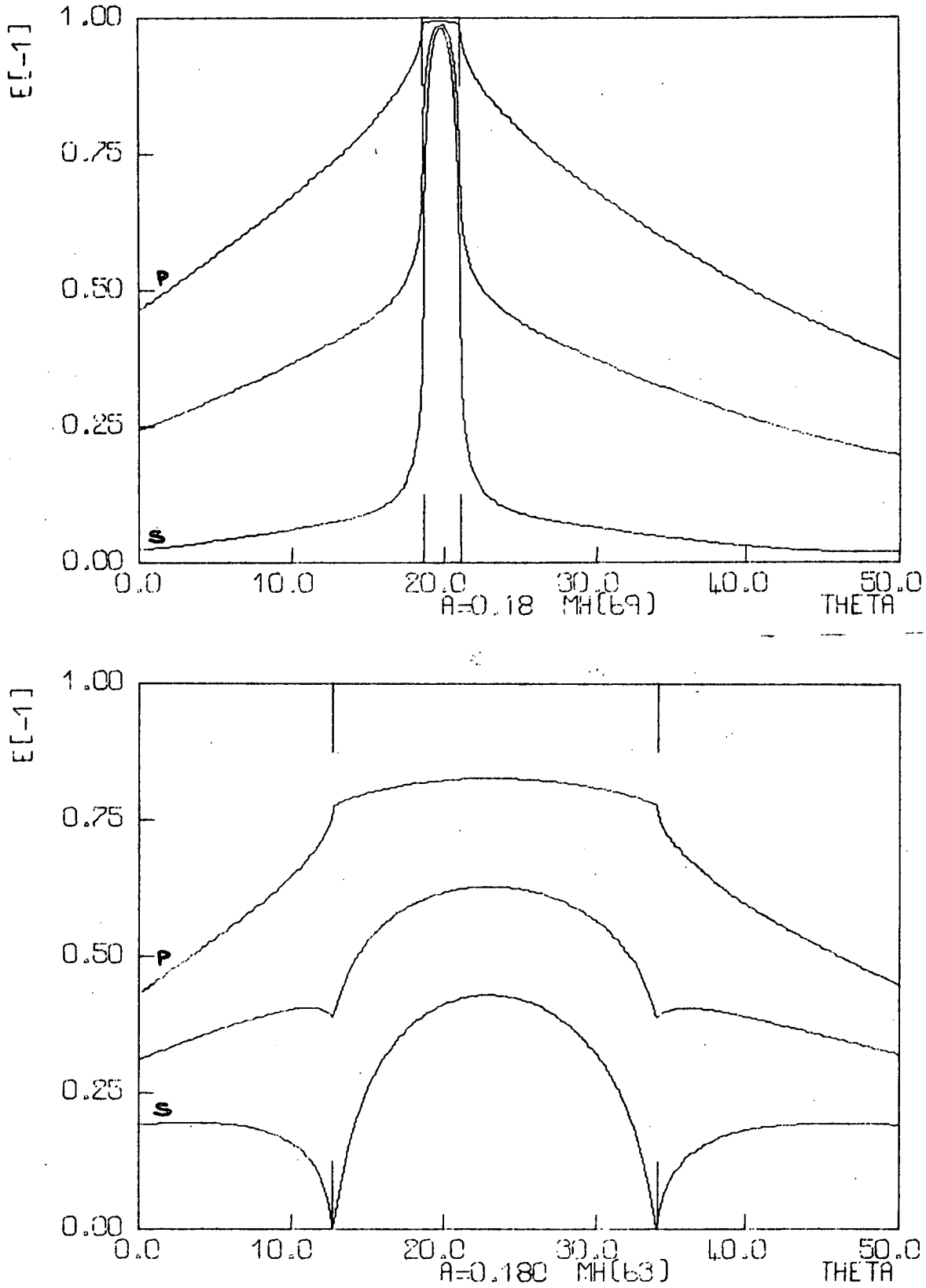


Figure (8.27). These theoretical efficiency curves refer to the behaviour of a grating having a normalized groove depth of 0.36 and the non-inverted profile (3).

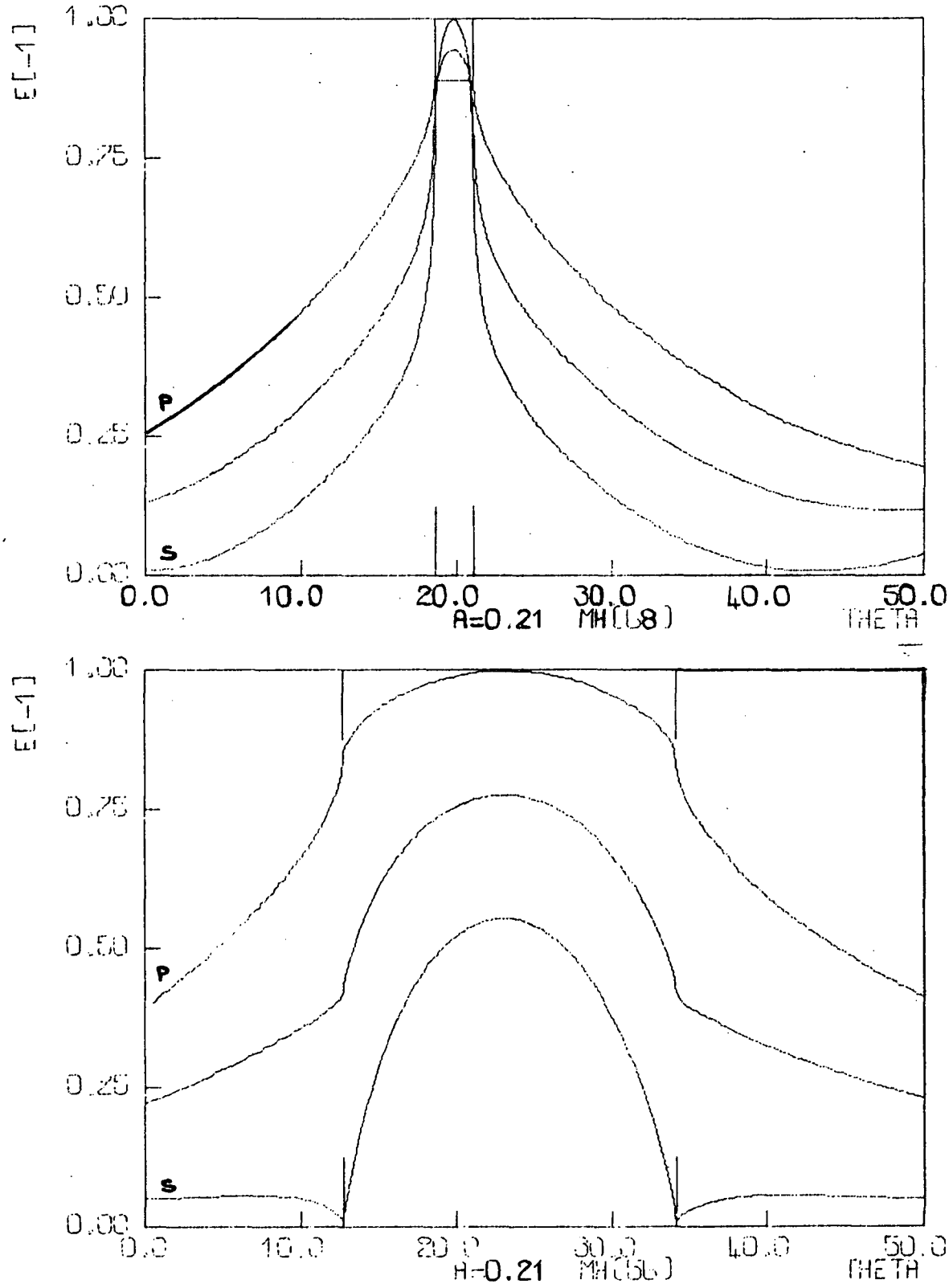


Figure (8.28). These theoretical efficiency curves refer to the behaviour of a grating having a normalized groove depth of 0.42 and a sinusoidal profile.

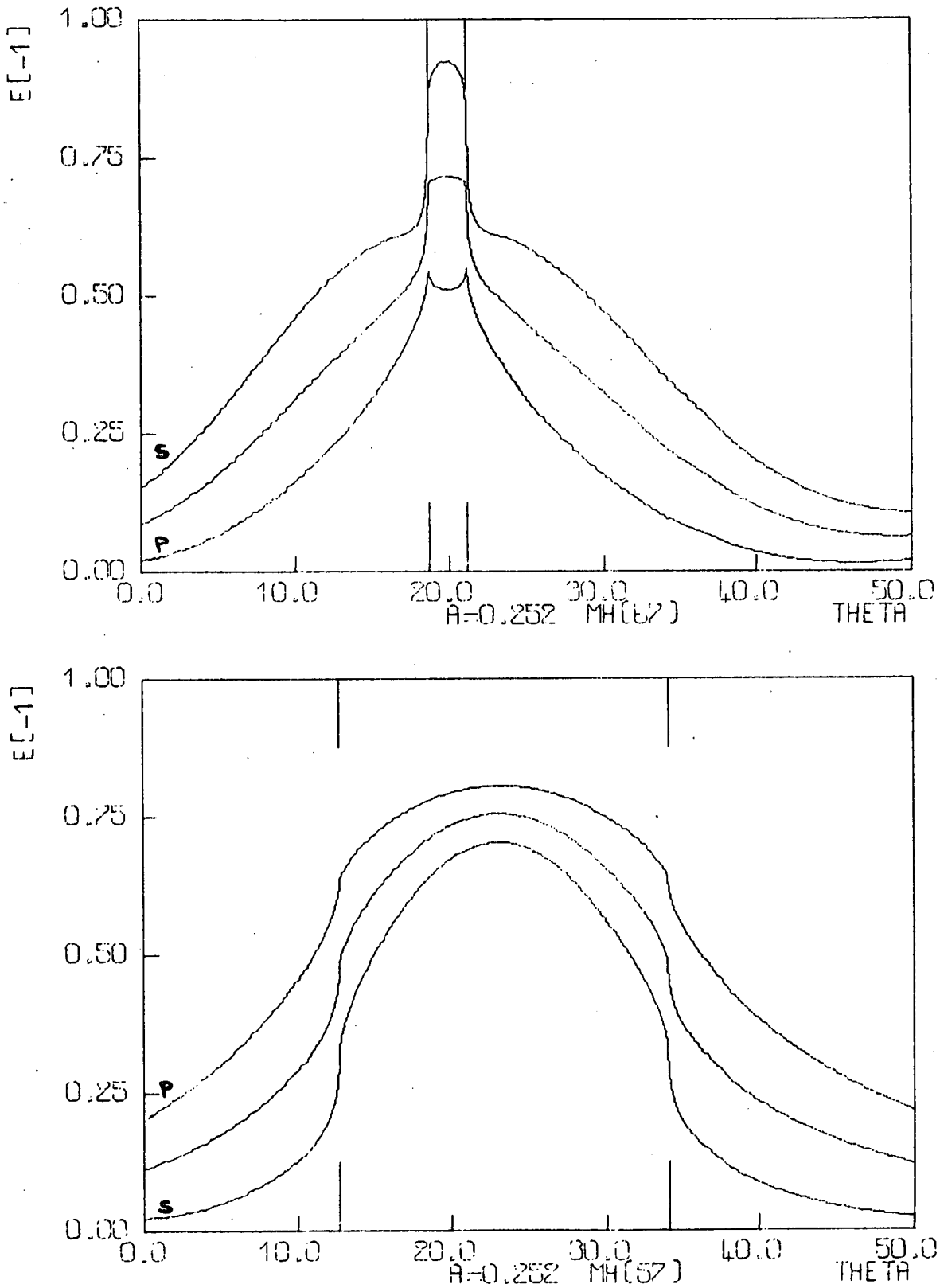


Figure (8.29). These theoretical efficiency curves refer to the behaviour of a grating having a normalized groove depth of 0.504 and the inverted profile (2).

particularly for shallower grooves. Away from this wavelength, with well-separated anomalies, the agreement between theory and experiment is quite satisfactory, bearing in mind the finite surface conductivity of the experimental gratings, and the possibility of profile variations across their surfaces.

It is reassuring that the agreement between theory and experiment starts to improve as the normalized groove depth approaches the values which we have seen in Chapter 6 provide the best blazewidths in unpolarized light for the sinusoidal profile.

The important effects which can result from profile distortion in the regions around Wood anomalies are quite evident. Clearly, it is desirable that the extent of the distortion actually present on the experimental profiles be determined, as a function of groove depth, if this is possible. It would seem from the above comparisons of theory and experiment that the distortion tends to increase as the grooves become shallower.

8.5. PROFILE DISTORTION AND EFFICIENCY CURVES

It was seen in the previous section that, for a grating with a normalized groove depth of 0.36, the use of the inverted distorted profiles results in a widening of the P polarization blaze region. The effect of the distortion for this polarization is to give an efficiency curve rather similar to one for a sinusoidal profile with grooves deeper than 0.36. Conversely, the use of non-inverted profiles worsens P polarization performance, by making the equivalent sinusoidal profile more shallow. Also, for the normalized groove depth of 0.36, the S polarization efficiency curve is not highly sensitive to inverted profile distortion.

It has been confirmed for groove depths other than 0.36 that grating performance in P polarized light can be improved through the use of inverted profile distortion, while the S polarization performance is not significantly impaired. Thus, the performance of the grating in unpolarized light can also be improved through the use of appropriate profile distortions.

Using the non-linear characteristics of Shipley AZ1350 photoresist, holographic gratings can be constructed whose profiles incorporate the non-inverted type of profile distortion. If these gratings are then replicated, their grooves will be inverted (8.12). Hence, the replica gratings will incorporate the desired inverted type of profile distortion.

Blaze optimization studies have been performed on a range of profiles incorporating various amounts of inverted distortion. For each profile shape, the optimal groove depth has been found, from the point of view of achieving the largest possible blazewidth associated with a minimum efficiency level of 80% and unpolarized light. Efficiency curves for P, S and U polarized light for the best of the depth-optimized profiles are shown in figure (8.30). For this profile, the blazewidths for the 60%, 70%, 80% and 90% minimum efficiency levels and unpolarized light are given in Table (8.1), together with the corresponding values for the optimized sinusoidal profile. The great improvements in spectral performance at the higher efficiency levels which can be obtained through the use of appropriately distorted profiles are made quite evident by figure (8.30) and Table (8.1).

8.6. CONCLUSIONS

For triangular profile gratings, it has been seen that for P

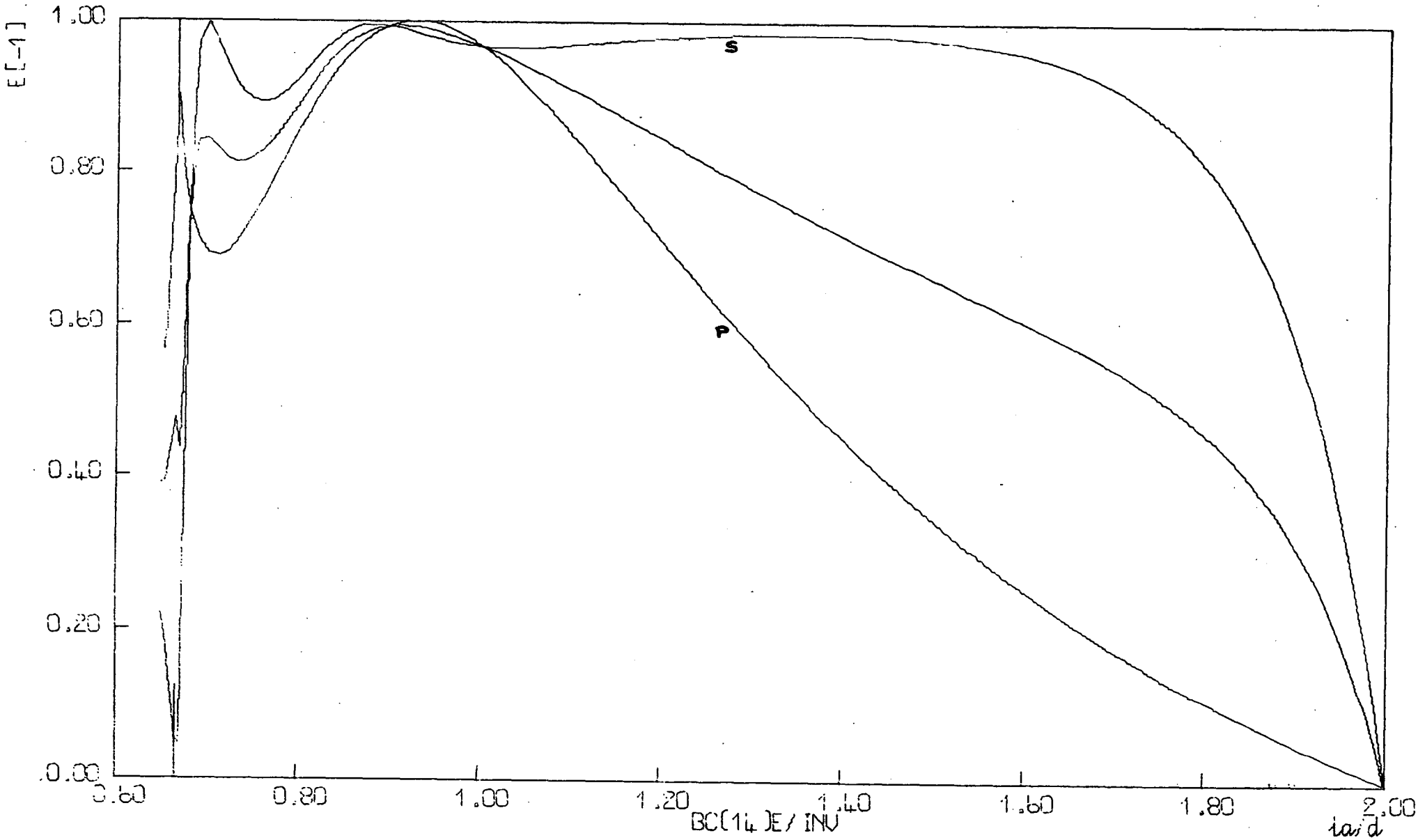


Figure (8.30). These theoretical efficiency curves correspond to a grating having an optimized inverted profile distortion, used in a Littrow mount in the order -1.

TABLE (8.1)

Normalized U polarization blazewidth values are given for optimized sinusoidal and distorted-sinusoidal profiles.

MINIMUM EFFICIENCY LEVEL	BLAZEWIDTH		% INCREASE
	SINE	DISTORTED SINE	
60%	0.83	0.93	13
70%	0.55	0.75	37
80%	0.33	0.58	78
90%	0.12	0.31	157

polarization the infinite conductivity formulism of Pavageau and Bousquet gives efficiency values which are in quantitative accord with experimental relative efficiency measurements. For S polarization, calculations and measurements are in qualitative agreement, with the theory tending to predict resonant spikes and minima which are sharper than those actually observed in the visible region.

For quasi-sinusoidal holographic gratings, it has been shown that again the theoretical predictions for P polarization are in satisfactory agreement with visible measurements. The situation is less satisfactory for S polarization, particularly when anomalies are formed close together.

It has been demonstrated that S polarization grating anomalies are quite sensitive to the effects of profile distortion. Thus, if their shape is to be measured in the visible region and compared with theoretical predictions, particular care must be taken to ensure that the profile of the experimental grating is accurately known. It may prove that further refinements of profile observation techniques are necessary before measurements and calculations for S polarization come into good agreement. The tolerances on the accuracy of the profile determination are less strict for P polarization, because of the lower sensitivity of anomaly shapes to the effects of profile deformation.

It is not possible at this stage to assess with certainty the relative contributions of finite surface conductivity and profile variation and distortion to the discrepancies between efficiency values calculated using an infinite-conductivity formulism and values measured in the visible region. The assessment will not be able to be made until the finite conductivity theory has been further developed, and then applied to give counterparts for experimental efficiency curves. However, based on experimental and theoretical data for triangular profile

gratings, E.G. Loewen (8.1) has made the following comment:

"In fact I think one can even judge the degree of departure from ideal groove profile by noting the degree of mismatch between theoretical curves and experiment. I do not believe that lack of perfect conductivity plays a significant role in practice, especially compared to imperfect groove shape."

The results of this chapter make clear the importance of achieving good control over the profile form of diffraction gratings. The rewards in terms of greatly improved spectral performance which will follow upon the achievement of such control are equally clear.

REFERENCES

(8.1) Loewen, E.G., 1973, personal communications.

(8.2) Deleuil, R., 1969, Optica Acta, 16, 23.

(8.3) Petit, R., 1966, Rev.d'Opt., 45, 353.

(8.4) Hutley, M.C., 1973, National Physical Laboratory, Division of Mechanical and Optical Metrology, Report MOM1.

(8.5) Hutley, M.C., 1973, paper submitted to Optica Acta.

(8.6) Palmer, C.H., 1952, J.opt.Soc.Amer., 42, 269.

(8.7) Palmer, C.H., 1956, J.opt.Soc.Amer., 46, 50.

(8.8) Stewart, J.E., and Gallaway, W.S., 1962, Appl.Optics, 1, 421.

(8.9) Wood, R.W., 1912, Phil.Mag., 23, 310.

(8.10) Strong, J., 1936, Phys.Rev., 49, 291.

(8.11) Beesley, M.J., and Castledine, J.G., 1970, Appl.Optics, 9, 2720.

(8.12) Stroke, G.W., 1967, in "Encyclopedia of Physics", 29, 732 (Berlin: Springer-Verlag).

APPENDIX I.

THE COMPUTATION OF GRATING EFFICIENCIES USING THE FORMULATION OF PAVAGEAU AND BOUSQUET.

In this appendix, various aspects of the programming of the theoretical formulism of Pavageau and Bousquet described in section five of Chapter 2 will be considered.

THE CALCULATION OF THE KERNEL FUNCTION - ANALYTICAL STUDY

We will now consider in detail the analytic properties of the kernel function N(x,t) for the case of S polarization, following the methods used by Pavageau and Bousquet (2.25) in their study of the P polarization kernel.

For S polarization,

$$N(x,t) = \sum_{p=-\infty}^{\infty} N_p = \sum_{p=-\infty}^{\infty} \left(s - \frac{\alpha_p}{\gamma_p} f'(t)\right) \exp j[pK(t-x) - k\gamma_p |f(t)-f(x)|], \quad (1)$$

where

$$\alpha_p = \sin i + \frac{pK}{k} , \quad (2)$$

$$\gamma_p = \sqrt{1-\alpha_p^2} \quad \text{if } \alpha_p^2 \leq 1$$
$$= -j\sqrt{\alpha_p^2-1} \quad \text{if } \alpha_p^2 > 1 \quad (3)$$

and

$$s = \text{sign}(f(x) - f(t)) . \quad (4)$$

The kernel function is thus the sum of an infinite series of oscillatory terms. If $f(t) \neq f(x)$, the convergence of the series is guaranteed by equation (3), which ensures that it is exponentially attenuated for

large $|p|$. The behaviour of $N(x,t)$ for values of x and t such that $f(x) = f(t)$ needs special consideration.

Let
$$N(x,t) = T(x,t) + R(x,t) , \tag{5}$$
where
$$T(x,t) = \sum_{p=-(L-1)}^{L-1} N_p \tag{6}$$
and
$$R(x,t) = \sum_{|p| \geq L} N_p . \tag{7}$$
Here L is a positive integer.

Consider first the behaviour of the truncated kernel T for values of t such that $f(t)$ is close to $f(x)$. Let $t=x'$ denote any solution of the equation $f(t) = f(x)$ such that $x' \neq x$. Since T has only a finite number of terms, it is not difficult to determine its limiting values as t tends to x and x':
$$\lim_{t\to x} T(x,t) = s(2L-1) - f'(x) \sum_{p=-(L-1)}^{L-1} \frac{\alpha_p}{\gamma_p} , \tag{8}$$
and
$$\lim_{t\to x'} T(x,t) = \frac{s \sin[(L-\frac{1}{2})K(x'-x)]}{\sin[\frac{K}{2}(x'-x)]} - f'(x') \sum_{p=-(L-1)}^{L-1} \frac{\alpha_p}{\gamma_p} \exp[jpK(x'-x)]. \tag{9}$$

Consider next the behaviour of the remainder R, firstly for t close to x. The integer L is chosen to be large enough so that the following expansions for $|p| \geq L$ are useful:
$$\gamma_p \simeq -\frac{j|p|K}{k}\left[1 + \frac{k\alpha_0}{pK} - \frac{k^2}{2p^2K^2} + O\left(\frac{1}{p^3}\right)\right] \tag{10}$$
$$\frac{\alpha_p}{\gamma_p} \simeq \frac{jp}{|p|}\left[1 + \frac{k^2}{2p^2K^2} + O\left(\frac{1}{p^3}\right)\right] . \tag{11}$$

Let

$$\xi = K(t - x) \tag{12}$$

and

$$\zeta = K[f(t) - f(x)] \, . \tag{13}$$

Then

$$R \simeq \sum_{|p|\geq L} \left(s - \frac{\alpha_p}{\gamma_p} f'(t)\right) \exp\left[jp\xi - |p\zeta|\left(1 + \frac{k\alpha_0}{pK} - \frac{k^2}{2p^2K^2}\right)\right]$$

$$\simeq \sum_{|p|\geq L} \left(s - \frac{\alpha_p}{\gamma_p} f'(t)\right) \exp\left[jp\xi - |p\zeta|\left(1 + \frac{k\alpha_0}{pK}\right)\right]$$

$$+ \frac{k^2|\zeta|}{2K^2} \sum_{|p|\geq L} \frac{1}{|p|} \left(s - \frac{\alpha_p}{\gamma_p} f'(t)\right) \exp\left[jp\xi - |p\zeta|\left(1 + \frac{k\alpha_0}{pK}\right)\right] + \ldots \tag{14}$$

Now

$$\sum_{p\geq L}^{\infty} \frac{1}{p} \exp\left[jp\xi - p|\zeta|\left(1 + \frac{k\alpha_0}{pK}\right)\right] \leq \sum_{p\geq L}^{\infty} \frac{\exp(-p|\zeta|)}{p} \, ,$$

and a consideration of the Taylor series of $\log(1-x)$ shows that the right-hand series diverges as $\log|\zeta|$ when $|\zeta| \to 0$. Thus, the second term in equation (14) tends to zero as $|\zeta| \log|\zeta|$ when $|\zeta| \to 0$, and so can be ignored. In the first term of equation (14), put

$$\frac{\alpha_p}{\gamma_p} = \left(\frac{\alpha_p}{\gamma_p} - \frac{jp}{|p|}\right) + \frac{jp}{|p|} \, .$$

The quantity in the brackets is of order p^{-2} and so gives rise to an absolutely and uniformly convergent series, with the limit

$$\sum_{|p|\geq L} \frac{\alpha_p}{\gamma_p} \, .$$

Thus,

$$R \simeq - f'(x) \sum_{|p| \geq L} \frac{\alpha_p}{\gamma_p} + \sum_{|p| \geq L} \left[s - \frac{jp}{|p|} f'(t) \right] \left(1 - \frac{k\alpha_0}{pK} |p\zeta|\right) \exp(jp\xi - |p\zeta|) \quad . \qquad (15)$$

Let

$$J = \sum_{p=L}^{\infty} \exp[p(j\xi - |\zeta|)] = \frac{\exp[L(j\xi - |\zeta|)]}{1 - \exp(j\xi - |\zeta|)}$$

$$= J_r + jJ_i \quad . \qquad (16)$$

Then

$$\sum_{p=-L}^{-\infty} \exp[p(j\xi + |\zeta|)] = J_r - jJ_i \quad . \qquad (17)$$

Thus,

$$R \simeq -f'(x) \sum_{|p| \geq L} \frac{\alpha_p}{\gamma_p} + 2\left[s + \frac{jf'(t)k\alpha_0|\zeta|}{K} \right] J_r - 2j\left[\frac{sk\alpha_0|\zeta|}{K} + jf'(t) \right] J_i \quad . \qquad (18)$$

For small ξ and ζ:

$$J \simeq \frac{1}{|\zeta| - j\xi} - (L - \tfrac{1}{2}) + O(\xi, \zeta) \quad ,$$

where $$|\zeta| = |\xi f'(x) + \frac{\xi^2 f''(x)}{2K} + \ldots| = -s\zeta$$

$$\therefore \lim_{t \to x} R = -f'(x) \sum_{|p| \geq L} \frac{\alpha_p}{\gamma_p} + \lim_{t \to x} \left\{ 2\left[s + \frac{j\, f'(t)k\alpha_0}{K} (-s\, f'(x)\xi) \right] \cdot \left[\frac{-s\, f'(x)}{\xi(1+f'(x)^2)} - \frac{s\, f''(x)}{2K(1+f'(x)^2)} - (L - \tfrac{1}{2}) + \frac{s\, f'(x)^2\, f''(x)}{K(1+ f'(x)^2)^2} \right] + \left[2f'(t) + \frac{2j\, k\alpha_0 f'(x)\xi}{K} \right] \left[\frac{1}{\xi(1+f'(x)^2)} - \frac{f'(x)\, f''(x)}{K(1+ f'(x)^2)^2} \right] \right\} \qquad (19)$$

$$\therefore \lim_{t\to x} R = - f'(x) \sum_{|p|\geq L} \frac{\alpha_p}{\gamma_p} - s(2L-1) + \frac{2j\, k\alpha_0 f'(x)}{K} - \frac{f''(x)}{K(1+f'(x)^2)}$$

$$+ \lim_{t\to x} \left\{ \frac{2(f'(t)-f'(x))}{\xi(1+f'(x)^2)} \right\} \qquad (20)$$

The final expression for this limit is

$$\lim_{t\to x} R(x,t) = -f'(x) \sum_{|p|\geq L} \frac{\alpha_p}{\gamma_p} - s(2L-1) + \frac{2j\, k\alpha_0 f'(x)}{K} + \frac{f''(x)}{K(1+f'(x)^2)} \; . \qquad (21)$$

Combining equations (8) and (21), we find that

$$\lim_{t\to x} N(x,t) = - f'(x) \sum_{p=-\infty}^{\infty} \frac{\alpha_p}{\gamma_p} + \frac{f''(x)}{K(1+f'(x)^2)} + \frac{2j\, k\alpha_0 f'(x)}{K} \; . \qquad (22)$$

Because of the presence of the sign quantity s in equation (1), $N(x,x)$ is undefined. However, equation (22) gives meaning to this value of the kernel function.

Consider next the behaviour of R for t close to x'. In this case, ξ does not tend to zero with ζ. Thus, the expression corresponding to equation (15) is

$$R \simeq - f'(x') \sum_{|p|\geq L} \left(\frac{\alpha_p}{\gamma_p} - \frac{jp}{|p|} \right) \exp(jp\xi) + \sum_{|p|\geq L} \left[s - \frac{jp}{|p|} f'(t) \right]$$

$$\left(1 - \frac{k\alpha_0}{pK} |p\zeta| \right) \exp(jp\xi - |p\zeta|) \; . \qquad (23)$$

Let

$$\left(\frac{\alpha_p}{\gamma_p} \right)' = \frac{\alpha_p}{\gamma_p} \quad \text{for } |p| < L$$

$$= \left(\frac{\alpha_p}{\gamma_p} - \frac{jp}{|p|} \right) \quad \text{for } |p| \geq L \qquad (24)$$

In this case, the function J defined in equation (16) remains bounded as t approaches x'. We have:

$$J \simeq -[\sin(L - \tfrac{1}{2})\xi - j\cos(L - \tfrac{1}{2})\xi]/2\sin(\tfrac{\xi}{2}) . \quad (25)$$

Hence,

$$\lim_{t\to x'} R = \frac{-s\sin(L - \frac{1}{2})K(x'-x) + f'(x')\cos(L - \frac{1}{2})K(x'-x)}{\sin\frac{K}{2}(x'-x)}$$

$$- f'(x') \sum_{|p|\geq L}' (\frac{\alpha_p}{\gamma_p}) \exp[jpK(x'-x)] . \quad (26)$$

Combining equations (9) and (26), we find that

$$\lim_{t\to x'} N(x,t) = f'(x')\left\{\frac{\cos(L - \frac{1}{2})K(x'-x)}{\sin\frac{K}{2}(x'-x)} - \sum_{p=-\infty}^{\infty}{}' (\frac{\alpha_p}{\gamma_p}) \exp[jp\,K(x'-x)]\right\} \quad (27)$$

Equations (22) and (27) define the values of N(x,t) in the two neighbourhoods where the right-hand side of equation (1) loses meaning. Elsewhere, the problem of the evaluation of the kernel function is one of summing a convergent infinite series.

THE CALCULATION OF THE KERNEL FUNCTION AT ORDINARY POINTS

Consider the pth term of the infinite series (1) for the kernel function, at points such that ζ is not close to zero. We will derive an asymptotic formula for N_p, based on the assumption that terms of order $1/p^4$ or larger are negligible.

Calculating an extra term for equations (10) and (11) gives

$$\gamma_p = -\frac{j|p|K}{k}\left[1 + \frac{k\alpha_0}{pK} - \frac{k^2}{2p^2K^2} + \frac{k^3\alpha_0}{2p^3K^3} + O(\frac{1}{p^4})\right] \quad (28)$$

and

$$\frac{\alpha_p}{\gamma_p} = \frac{jp}{|p|}\left[1 + \frac{k^2}{2p^2K^2} - \frac{k^3\alpha_0}{K^3p^3} + O\left(\frac{1}{p^4}\right)\right]. \qquad (29)$$

$$\therefore \quad N_p = \left(s - \frac{\alpha_p}{\gamma_p} f'(t)\right) \exp\left(jp\xi - \frac{jk\,\gamma_p|\zeta|}{K}\right)$$

$$= \left(s - \frac{jp}{|p|} f'(t) - \frac{jk^2}{2K^2p|p|} f'(t) + \frac{j\,k^3\alpha_0 f'(t)}{K^3p^2|p|}\right)\exp(jp\xi - |p\zeta|).$$

$$\left[\exp\left(-\frac{k\alpha_0}{K}|\zeta|\right)\right]^{|p|/p} \exp\left[\frac{|\zeta|k^2}{2p^2K^2}|p|\left(1 - \frac{k\alpha_0}{pK}\right.\right.$$

$$\left.\left. + \frac{k^2}{4p^2K^2}(1 + 4\alpha_0^2)\right)\right] \qquad (30)$$

For simplicity, we will consider the expression (30) separately for positive and negative values of p. Let N_p^+ denote the value of N_p for positive p. For $p < 0$, we will replace p by (-p) in order to calculate N_p^-. We will make the following substitutions:

$$A = s - j\,f'(t), \qquad A^* = s + j\,f'(t), \qquad (31)$$

$$B = \frac{k^2 f'(t)}{2K^2}, \quad (32) \qquad C = \frac{k^2|\zeta|}{2K^2}, \qquad (33)$$

$$D = \exp\left(-\frac{k\alpha_0|\zeta|}{K}\right), \quad (34) \text{ and } \quad E = \frac{k^3\alpha_0 f'(t)}{K^3}. \qquad (35)$$

$$\therefore N_p^+ = \left(A - \frac{jB}{p^2} + \frac{jE}{p^3}\right) D \exp(jp\xi - p|\zeta|) \exp\left[\frac{C}{p}\left(1 - \frac{k\alpha_0}{pK}\right.\right.$$

$$\left.\left. + \frac{k^2}{4p^2K^2}(1 + 4\alpha_0^2)\right)\right]$$

$$= \left(A - \frac{jB}{p^2} + \frac{jE}{p^3}\right)\left[1 + \frac{C}{p} + \left(\frac{C^2}{2} - \frac{Ck\alpha_0}{K}\right)\frac{1}{p^2} + \frac{1}{p^3}\left(\frac{C^3}{6} - \frac{C^2k\alpha_0}{K}\right.\right.$$

$$\left.\left. + \frac{Ck^2}{4K^2}(1 + 4\alpha_0^2)\right)\right] D \exp\left[(j\xi - |\zeta|)p\right]. \qquad (36)$$

Let $$z = |\zeta| - j\xi \ , \quad z^* = |\zeta| + j\xi \ . \tag{37}$$

Then, if the first two terms in equation (36) are multiplied together and terms of $O(\frac{1}{p^4})$ are neglected, we obtain

$$N_p^+ = \Big[A + \frac{AC}{p} + \frac{1}{p^2}\left(\frac{AC^2}{2} - \frac{AC\, k\alpha_0}{K} - jB \right)$$

$$+ \frac{1}{p^3}\left(\frac{AC^3}{6} + \frac{ACk^2}{4K^2}(1+4\alpha_0{}^2) - \frac{AC^2 k\alpha_0}{K} - jBC + jE \right)\Big] D \exp(-pz). \tag{38}$$

In a similar manner, the corresponding expression for N_p^- is obtained:

$$N_p^- = \Big[A^* + \frac{A^*C}{p} + \frac{1}{p^2}\left(\frac{A^*C^2}{2} + \frac{A^*Ck\alpha_0}{K} + jB \right)$$

$$+ \frac{1}{p^3}\left(\frac{A^*C^3}{6} + \frac{A^*Ck^2}{4K^2}(1+4\alpha_0{}^2) + \frac{A^*C^2 k\alpha_0}{K} + jBC + jE \right)\Big] \frac{\exp(-pz^*)}{D} \ . \tag{39}$$

Now,

$$R = \sum_{p=L}^{\infty} (N_p^+ + N_p^-) \ . \tag{40}$$

Let

$$\sum_{p=L}^{\infty} \frac{\exp(-pz)}{p^n} = S_{n,L}(z) \ . \tag{41}$$

Then

$$\sum_{p=L}^{\infty} \frac{\exp(-pz^*)}{p^n} = S_{n,L}^*(z) = S_{n,L}(z^*) \ . \tag{42}$$

Combining equations (38) and (39), we obtain the desired expression for the remainder of the kernel function:

$$R = AD\, S_{0,L}(z) + \frac{A^*}{D} S^*_{0,L}(z) + ACD\, S_{1,L}(z) + \frac{A^*C}{D} S^*_{1,L}(z)$$

$$+ D\left(\frac{AC^2}{2} - \frac{AC\, k\alpha_0}{K} - jB\right)S_{2,L}(z) + \frac{1}{D}\left(\frac{A^*C^2}{2} + \frac{A^*C\, k\alpha_0}{K} + jB\right)S^*_{2,L}(z)$$

$$+ D\left(\frac{AC^3}{6} + \frac{ACk^2}{4K^2}(1+4\alpha_0{}^2) - \frac{AC^2k\alpha_0}{K} - jBC + jE\right)S_{3,L}(z)$$

$$+ \frac{1}{D}\left(\frac{A^*C^3}{6} + \frac{A^*Ck^2}{4K^2}(1+4\alpha_0{}^2) + \frac{A^*C^2k\alpha_0}{K} + jBC + jE\right)S^*_{3,L}(z) \quad . \quad (43)$$

The truncated kernel is

$$T = \sum_{p=-(L-1)}^{L-1} \left(s - \frac{\alpha_p}{\gamma_p} f'(t)\right) \exp j\left[p\xi - \frac{k\gamma_p}{K}|\zeta|\right] \quad . \quad (44)$$

This can be evaluated by direct summation.

Thus, the problem of the numerical evaluation of the kernel function is reduced to one of finding the values of $S_{n,L}(z)$ for $n = 0,1,2$ and 3.

THE SUMMATION OF SOME INFINITE SERIES

For $n = 0$, $S_{n,L}(z)$ reduces to the sum of a geometric progression:

$$S_{0,L}(z) = \sum_{p=L}^{\infty} \exp(-pz) = \frac{\exp(-Lz)}{1-\exp(-z)} \quad . \quad (45)$$

For $n = 1$, $S_{n,L}(z)$ can also be evaluated analytically (2.26). We have:

$$S_{1,L}(z) = \sum_{p=L}^{\infty} \frac{\cos(p\xi)\exp(-p|\zeta|)}{p} + j \sum_{p=L}^{\infty} \frac{\sin(p\xi)\exp(-p|\zeta|)}{p} \quad . \quad (46)$$

Let

$$\theta = \xi, \quad q = |\zeta|/\xi \quad . \quad (47)$$

Then:

$$S_{1,L}(z) = C + iS - \sum_{p=1}^{L-1} \frac{\exp(-zp)}{p} \quad , \tag{48}$$

where

$$C = -\log_n \sqrt{1 - 2e^{-q\theta}\cos\theta + e^{-2q\theta}} \tag{49}$$

and

$$S = \arctan\left(\frac{\sin\theta}{e^{q\theta} - \cos\theta}\right) \quad . \tag{50}$$

The series for $S_{n,L}(z)$ converges at the same rate or faster than the sum of the geometrical progression $\exp(-p|\zeta|)$. Hence, if $|\zeta|$ is sufficiently large, only a small number of terms need to be added directly to obtain a good approximation for the sum of the series. For $|\zeta| \geq 1$, direct summation can be used to evaluate $S_{n,L}(z)$ for $n = 1,2$ and 3.

For $|\zeta| < 1$, the Euler-Maclaurin summation formula can be used to evaluate $S_{2,L}(z)$ and $S_{3,L}(z)$. From reference (2.27),

$$\sum_{p=L}^{\infty} \frac{\exp(-zp)}{p^n} \simeq \int_L^{\infty} \frac{\exp(-zx)}{x^n}\,dx + \frac{\exp(-zL)}{2L^n} - \frac{1}{12} f'(0)$$

$$+ \frac{1}{720} f^{III}(0) - \frac{1}{30,240} f^{V}(0) + \frac{1}{1,209,600} f^{VII}(0). \tag{51}$$

Let

$$F = e^{-zL} \quad . \tag{52}$$

Then

$$f'(0) = -\frac{F}{L^n}\left(z + \frac{n}{L}\right) \quad , \tag{53}$$

$$f^{III}(0) = -\frac{F}{L^n}\left(z^3 + \frac{3nz^2}{L} + \frac{3n(n+1)z}{L^2} + \frac{n(n+1)(n+2)}{L^3}\right), \tag{54}$$

$$f^{V}(0) = -\frac{F}{L^n}\left(z^5 + \frac{5nz^4}{L} + \frac{10n(n+1)z^3}{L^2} + \frac{10n(n+1)(n+2)z^2}{L^3}\right.$$

$$\left. + \frac{5n(n+1)(n+2)(n+3)z}{L^4} + \frac{n(n+1)(n+2)(n+3)(n+4)}{L^5}\right), \qquad (55)$$

and

$$f^{VII}(0) = -\frac{F}{L^n}\left(z^7 + \frac{7nz^6}{L} + \frac{21n(n+1)z^5}{L^2} + \frac{35n(n+1)(n+2)z^4}{L^3}\right.$$

$$+ \frac{35n(n+1)(n+2)(n+3)z^3}{L^4} + \frac{21n(n+1)(n+2)(n+3)(n+4)z^2}{L^5}$$

$$\left. + \frac{7n(n+1)(n+2)(n+3)(n+4)(n+5)z}{L^6} + \frac{n(n+1)(n+2)(n+3)(n+4)(n+5)(n+6)}{L^7}\right). \qquad (56)$$

Also,

$$\int_L^\infty \frac{\exp(-zx)dx}{x^n} = \frac{w(zL,n)e^{-zL}}{L^n z} \quad . \qquad (57)$$

The function w(zL,n) can be evaluated numerically using an algorithm based on a continued fraction representation (2.28).

The formulae of this section and its precursor enable the calculation of the kernel function at all points for which f(t) is not close to f(x). At other points, equations (22) and (27) are used to evaluate the kernel.

OTHER DETAILS OF THE PROCESS OF EFFICIENCY CALCULATION

Up to this point, we have concentrated on the evaluation of the S polarization kernel function. The calculation of the kernel function for P polarization poses no numerical problems other than those solved above for S polarization.

In order to calculate the efficiencies of the various

propagating orders, the current density function $\phi(x)$ must be determined. For both P and S polarizations, this is done by solving an integral equation of the form

$$\phi(x) = \phi_0(x) + \frac{1}{a} \int_0^a N(x,t)\, \phi(t)\, dt \ , \qquad (58)$$

where $\phi_0(x)$ is a known function.

Pavageau and Bousquet (2.24, 2.25) proposed that this equation be solved by an iterative method. $\phi_0(x)$ is regarded as the zeroth order approximation to ϕ. If $\phi_0(x)$ is substituted under the integral sign in equation (58), a first approximation $\phi^{(1)}(x)$ to $\phi(x)$ can be obtained. This process of substitution under the integral sign is repeated until the process converges to within a specified tolerance.

Maystre and Petit (2.29) suggested that, when the iterative method of solution fails, it can be replaced by a point-matching technique. This latter method is described further below. It has been used exclusively in the author's efficiency calculations using the formulism of Pavageau and Bousquet.

Let a quadrature rule of order m over the interval [0,a] use abscissae $x_1, x_2, \ldots, x_m$ and associated weights $\omega_1, \omega_2, \ldots, \omega_m$. Then:

$$\int_0^a N(x,t)\, \phi(t) dt \simeq \sum_{j=1}^{m} \omega_j\, N(x,x_j)\, \phi(x_j) \ . \qquad (59)$$

From equation (58),

$$\phi(x) - \sum_{j=1}^{m} \frac{\omega_j}{a} N(x,x_j)\, \phi(x_j) \simeq \phi_0(x) \ . \qquad (60)$$

Thus, for $i = 1,2,\ldots,m$,

$$\phi(x_i) - \sum_{j=1}^{m} \frac{\omega_j}{a} N(x_i, x_j)\, \phi(x_j) \simeq \phi_0(x_i) , \qquad (61)$$

or, equivalently,

$$\sum_{j=1}^{m} A_{i,j}\, \phi(x_j) \simeq \phi_0(x_i) , \qquad (62)$$

where

$$A_{i,j} = \delta_{i,j} - \frac{\omega_j}{a} N(x_i, x_j) , \qquad (63)$$

and $\delta_{i,j}$ is the Kronecker delta symbol.

The integral equation (59) has been reduced to an m×m inhomogeneous set of linear equations (62), with the unknowns being the complex quantities $\phi(x_1)$, $\phi(x_2)$,...,$\phi(x_m)$. If the real and imaginary parts of the $\phi(x_i)$ are taken as the unknowns, a 2m×2m set of linear equations with real coefficients can be derived from equation (62). This second set of equations can be solved by standard techniques.

In the Algol programs which have been written to calculate efficiency values using this formulation, the H.U.C.C. Library Procedure SOLVEQ is employed to solve the linear equations. SOLVEQ is based on a Gauss-Jordan reduction to diagonal form of the coefficients matrix, with partial pivoting. It has been found advantageous to machine-code part of this procedure; this change more than halved the run-time of the programs.

Once the values of $\phi(x_1)$, $\phi(x_2)$,...,$\phi(x_m)$ have been determined, the amplitudes of the various components of the diffracted field can be evaluated using the formula:

$$H_p \simeq \frac{1}{2a} \sum_{i=1}^{m} \phi(x_i)\, \omega_i \left(\frac{\alpha_p}{\gamma_p} f'(x_i) - 1 \right) \exp j[pKx_i + k\gamma_p f(x_i)]. \qquad (64)$$

The intensity associated with the order p is

$$I_p = |H_p|^2 \quad . \tag{65}$$

The efficiency associated with the propagating order p is

$$Z_p = I_p \frac{\gamma_p}{\gamma_0} \quad . \tag{66}$$

If the orders from LL to UL are propagating, the total diffracted energy is

$$D.E. = \sum_{p=LL}^{UL} Z_p \quad . \tag{67}$$

The conservation of energy dictates that D.E. should be one; its departure from unity furnishes the primary test on the accuracy of the calculations. A secondary test is provided by the evaluation of the amplitudes of the components of the diffracted field inside the grating surface, where the sum of incident and diffracted fields should be zero.

The numerical solution of the diffraction problem for P polarization follows along very similar lines to the S polarization solution described above. Only minor modifications distinguish the basic equations of the formulations for these two fundamental cases of polarization.

Programs have been written which calculate P and S polarization efficiencies together for triangular profiles (with and without land), and for sinusoidal profiles. A general profile program does not require an analytical formula for the profile to be specified - values of the profile function together with its first and second derivatives at the m integration abscissae are sufficient.

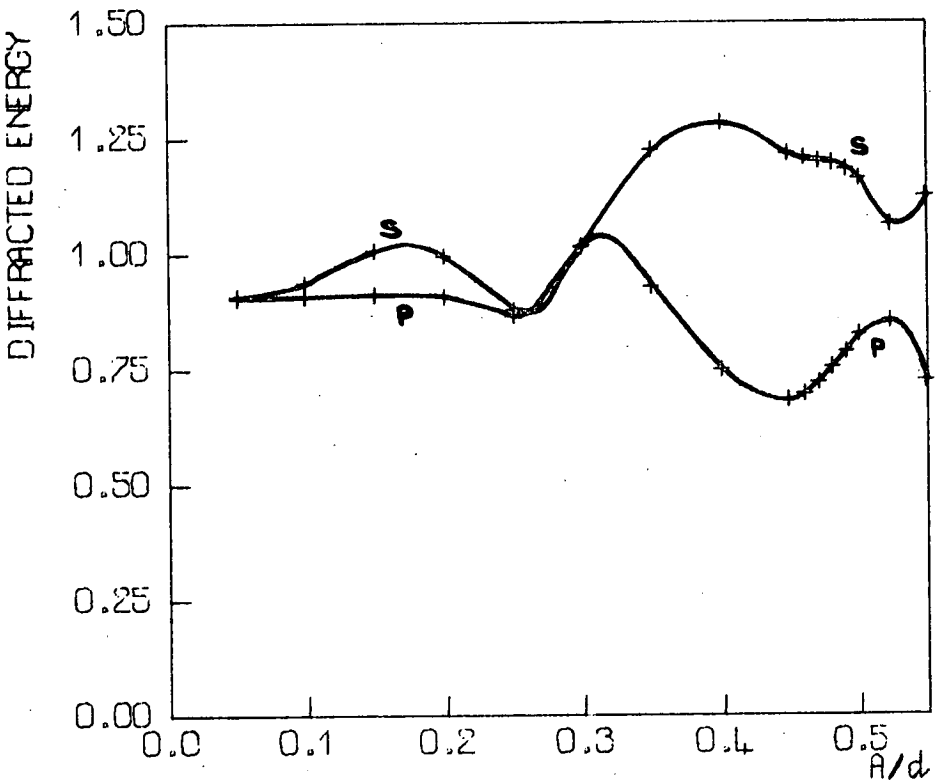


Figure 1. The variation of diffracted energy as a function of normalized groove depth is shown, for sinusoidal gratings used with normally incident light of wavelength $\lambda/d = 0.4368$. A composite twenty-point trapezoidal quadrature rule has been used.

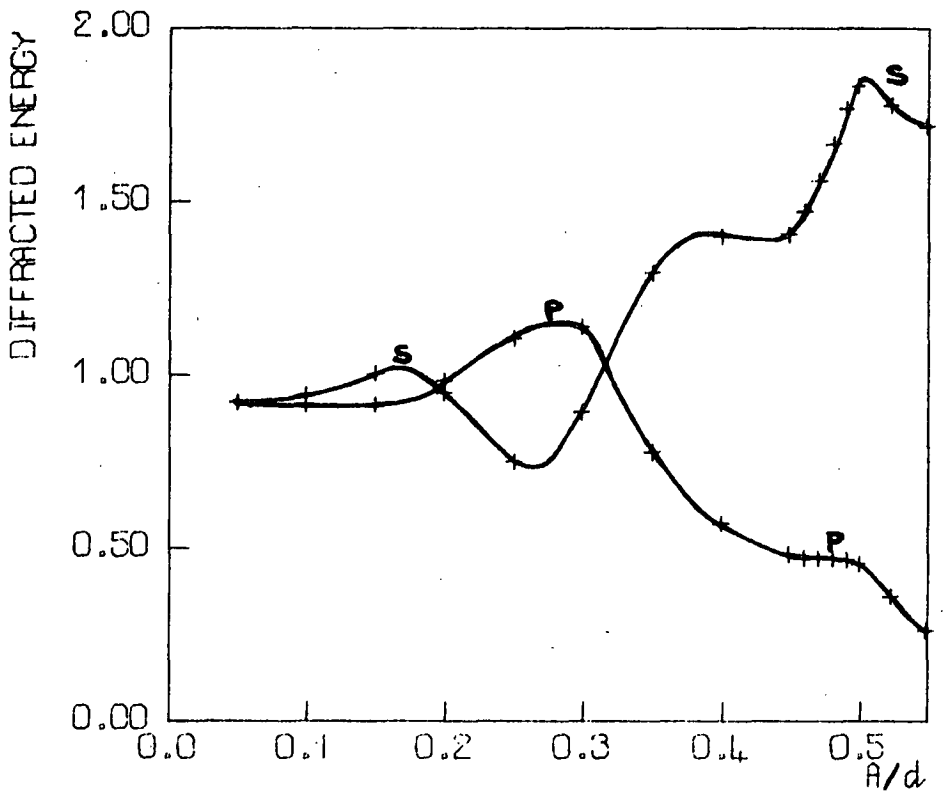


Figure 2. The diffracted energy curves correspond to the same data as those of figure 1, but have been obtained using a composite, twenty-point parabolic rule.

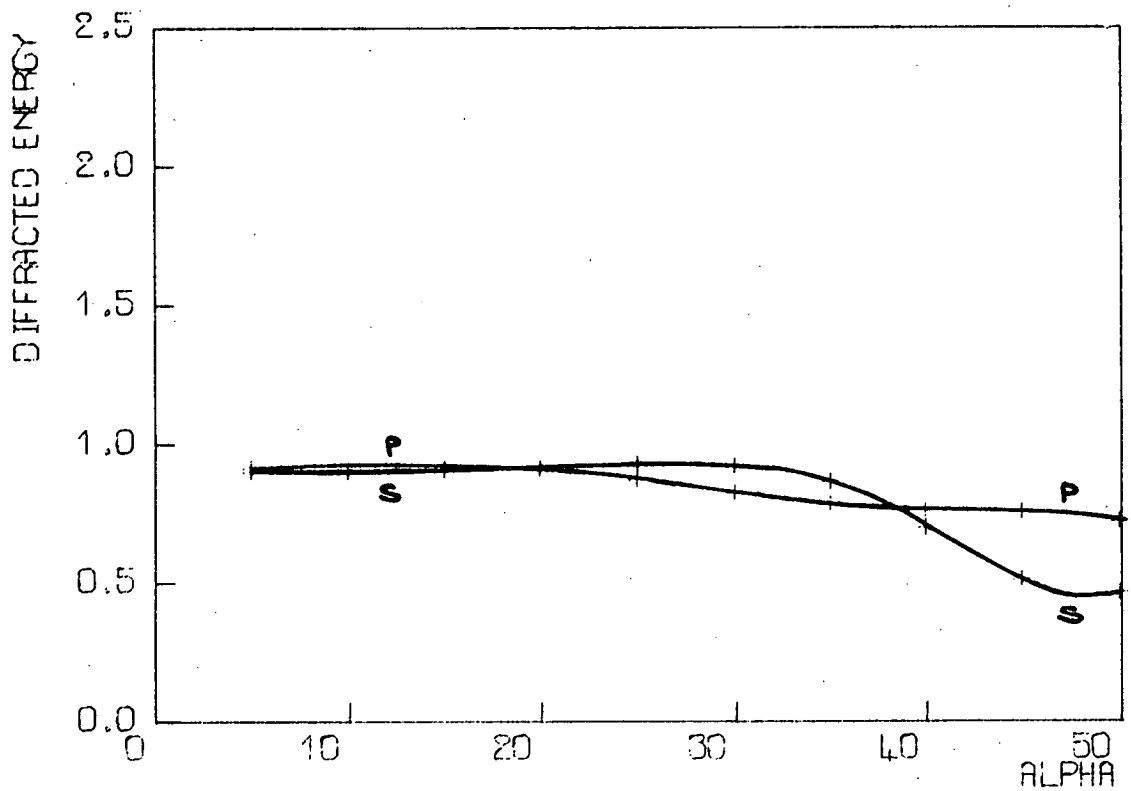


Figure 3. The variation of diffracted energy as a function of the blaze angle (α) in degrees is shown, for gratings having symmetrical triangular profiles, used with normally incident light of wavelength $\lambda/d = 0.4368$. A composite twenty-point trapezoidal quadrature rule has been used.

THE CHOICE OF THE QUADRATURE RULE

From Table (2.1), it can be seen how important it is in terms of computer time requirements to keep as low a value of the order of integration m as is consistent with accurate efficiency values. Hence, an investigation was undertaken to find the optimum quadrature rule for efficiency calculations for the two most important profile forms - the sinusoidal and triangular shapes.

Newton-Cotes quadrature rules of order m are exact for polynomials of degree m+1 or less. On the other hand, Gaussian quadrature rules of order m are exact for polynomials of degree up to 2m-1 (2.30). Since there is no constraint on our choice of the abscissae x_i, the unequal spacing of the points in Gaussian rules is not a disadvantage in this application. Thus, it would appear likely that Gaussian quadrature would be favoured over the use of Newton-Cotes integration formulae.

To confirm this, composite trapezoidal and parabolic twenty-point rules were used to obtain diffracted energy curves corresponding to those shown in Figure (2.17), which were based on a two-segment, twenty-point Gaussian rule. (The integration interval [0,a] is divided into two halves, and the abscissae and weights are calculated for ten-point quadratures over each half, in order to establish such a Gaussian rule. The calculations of the Gauss-Legendre weights and abscissae were performed using the procedure legendre written by D. Paget, of the Mathematics Department, University of Tasmania.)

Figures 1 and 2 show these curves giving diffracted energy as a function of normalized amplitude for sinusoidal gratings used with normally incident light of wavelength $\lambda/d = 0.4368$. The trapezoidal

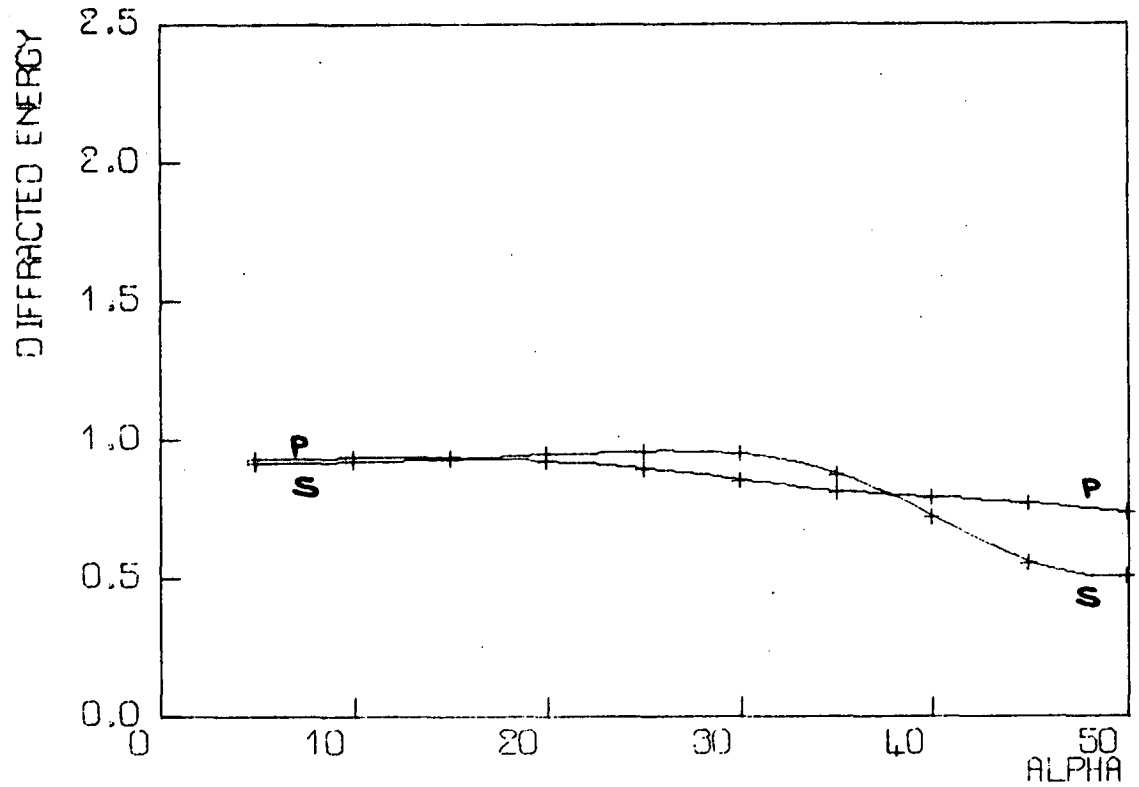


Figure 4. The diffracted energy curves correspond to the same data as those of figure 3, but have been obtained using a composite, twenty-point parabolic rule.

rule is seen to yield results of greater accuracy than does the parabolic rule. However, the superior accuracy which can be obtained by the use of Gaussian quadrature is quite evident.

Consider now the choice of the integration rule when the grating profile is triangular in form. It has been found beneficial to use two-segment Gaussian rules, with the groove apex corresponding to the point of segmentation. The choice of the number of points in each of the two segments is determined by the groove facet angles α and β. Twenty-point Gaussian rules have in general provided energy defects smaller than 1% in the efficiency calculations performed for triangular profiles. For m = 20 and a groove having a vertex angle of 90^{o}, the number of points (m1) placed on the longer groove facet is determined by the blaze angle α in the following fashion:

$$\text{for } \alpha \geq 30^{o},\ m1 = 11;$$
$$\text{for } 18^{o} \leq \alpha < 30^{o},\ m1 = 12;$$
$$\text{for } 13^{o} \leq \alpha < 18^{o},\ m1 = 13;$$
$$\text{for } \alpha < 13^{o},\ m1 = 14.$$

The diffracted energy curves of Figure (2.16) were calculated using a two-segment Gaussian twenty-point rule, for symmetrical triangular profiles used with normally incident light of wavelength $\lambda/d = 0.4368$. Corresponding curves obtained using twenty-point trapezoidal and parabolic rules are shown in Figures 3 and 4. Again, both these rules afford results of much worse accuracy than does a Gaussian formula of the same order.

Special care must be exercised in the choice of quadrature rule when the grating profile is composed of straight line elements which meet to form corners. At such corners, the profile derivative

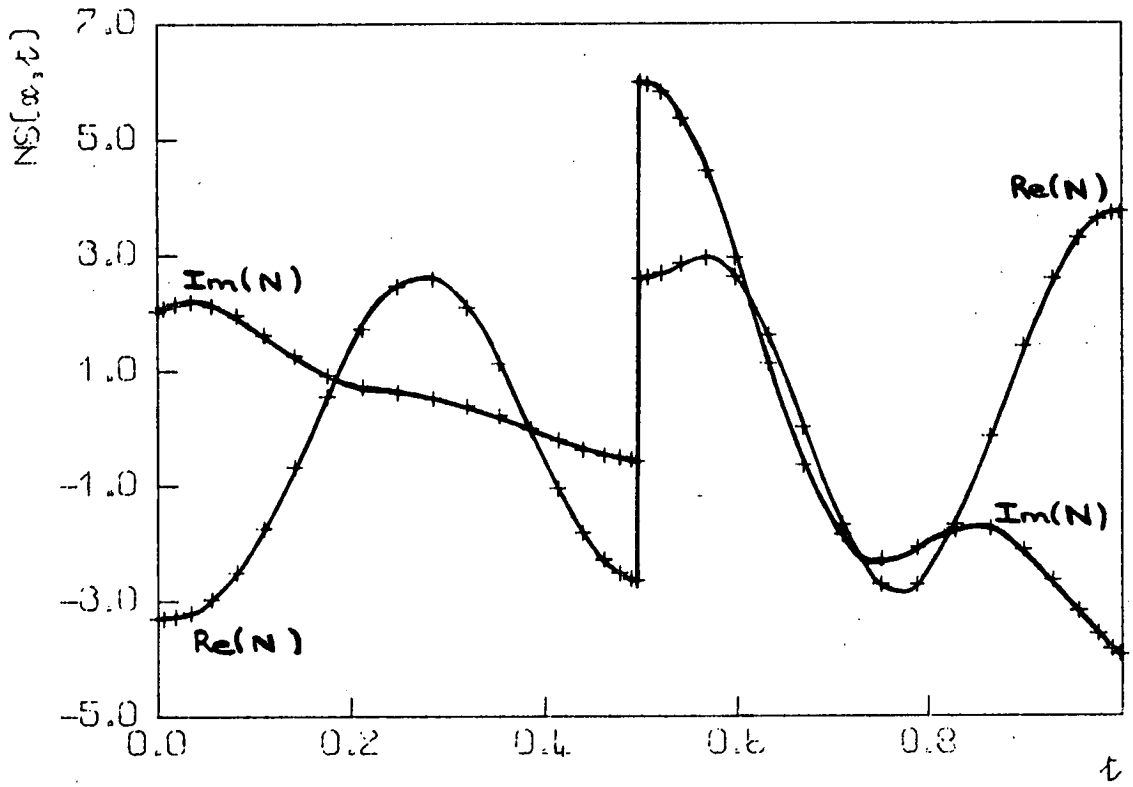


Figure 5. The variation of the S polarization kernel function NS(x,t) is shown as a function of t, for x = 0.3879. The calculation of NS was made for a grating having a symmetrical triangular profile with blaze angle 45^{o}, used with normally incident light of wavelength $\lambda/d = 0.45$. A two-segment, forty-point Gaussian quadrature rule was employed in the computation.

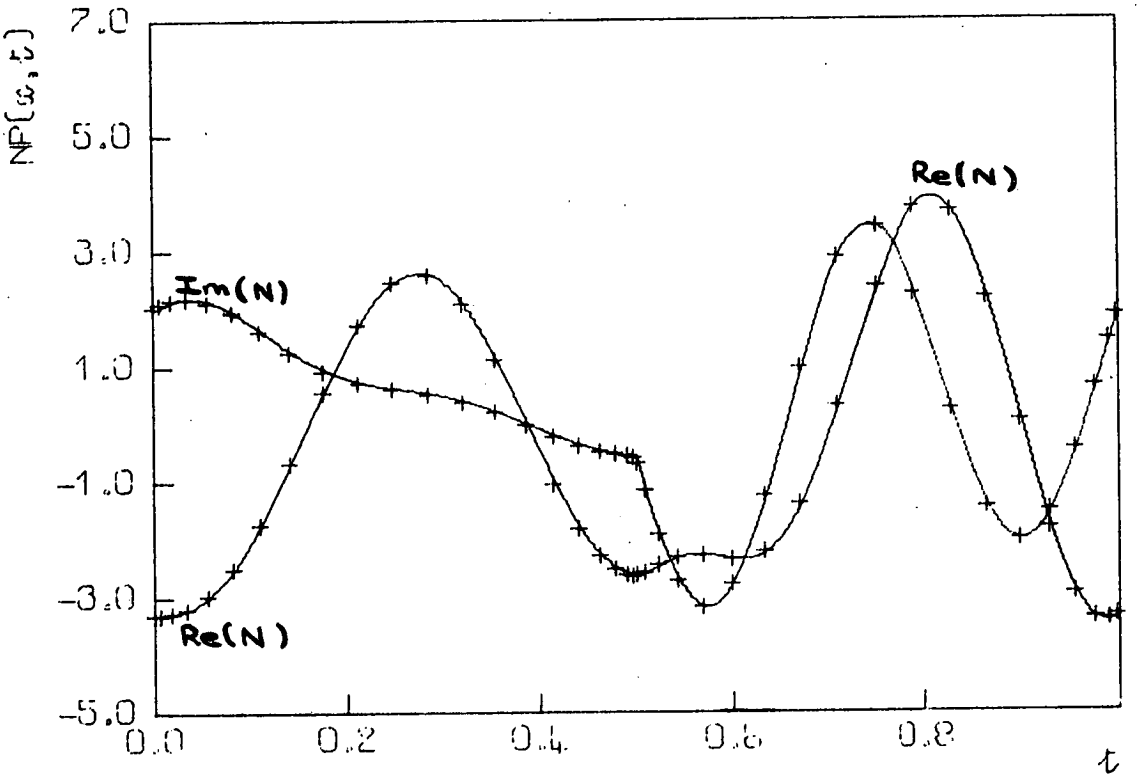


Figure 6. The curves correspond to the same data as those of figure 5, except that the behaviour of the P polarization kernel function NP(x,t) is shown.

f'(t) which occurs in the S polarization kernel function NS(x,t) is undefined. Thus, an integration rule must be chosen which does not incorporate a corner abscissa. However, a good concentration of integration abscissae is needed around the corner. This is evident for S polarization from Figure 5. Here the real and imaginary parts of NS(x,t) are plotted, for a grating having two groove facet angles of 45^{o}, used with normally incident light. The substantial step discontinuities of NS which occur at the groove corners (t = 0.0, 0.5, 1.0) make it necessary to have concentrations of integration abscissae there if integrals involving this function are to be evaluated accurately. For the P polarization kernel NP, Figure 6 shows that the discontinuities at groove corners occur only in the derivative of the function.

For P polarization, the need for concentration of abscissae around profile corners arises because of the behaviour of the current density function $\phi(P)$ rather than the behaviour of the kernel function. It has been shown (2.32) that at a corner having an included angle β, $\phi(P)$ will vary with the distance (r) from the corner as

$$r^{(\beta-\pi)/(2\pi-\beta)} \quad . \tag{68}$$

Thus, $\phi(P)$ will be singular at profile corners for which β is smaller than π, but not at corners for which β exceeds π. The S polarization current density $\phi(S)$ has no such singularities at either type of corner.

The occurrence of P polarization current density singularities can be seen in Figures 7 and 8, which correspond respectively to symmetrical profiles having facet angles of 15^{o} and 45^{o}. (Similar graphs have previously been given by Kalhor and Neureuther (2.33).) As might be expected from equation (68), the singularities in the real and

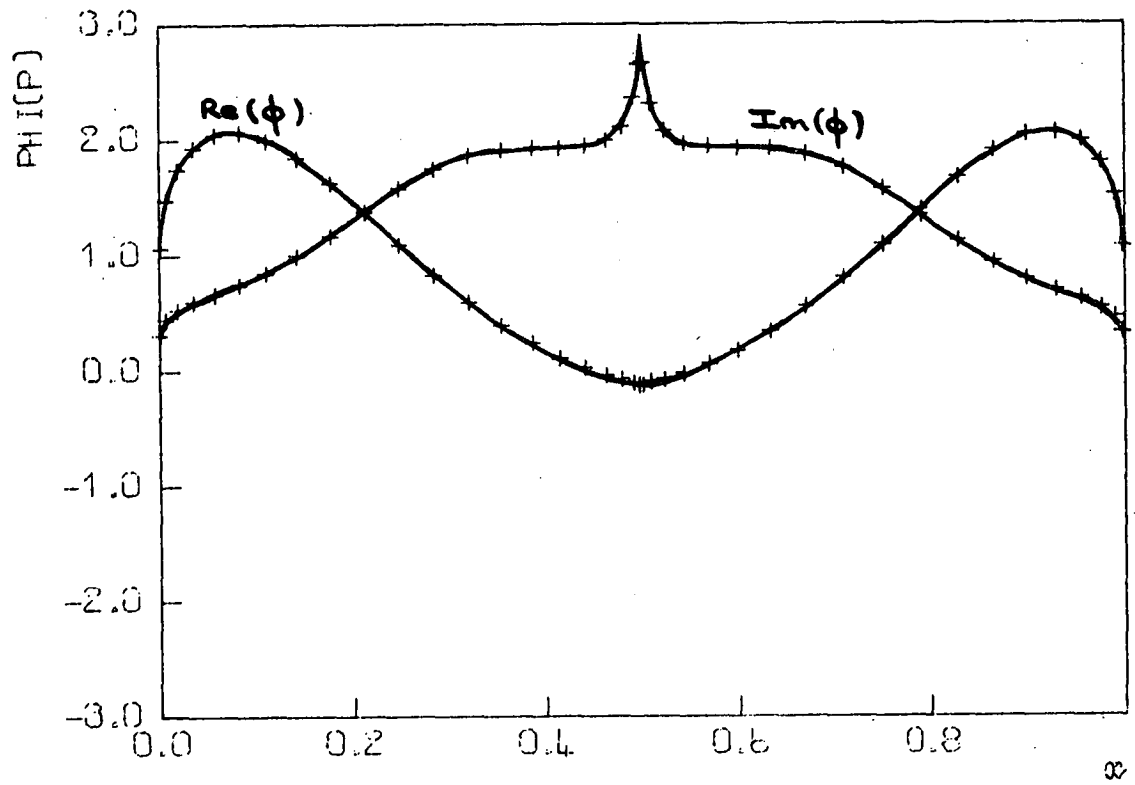


Figure 7. The variation of the P polarization current density function $\phi(P)(x)$ is shown, for a grating having a symmetrical triangular profile with blaze angle 15^{0}, used with normally incident light of wavelength $\lambda/d = 0.45$. A two-segment, forty-point Gaussian quadrature rule was employed in the computation.

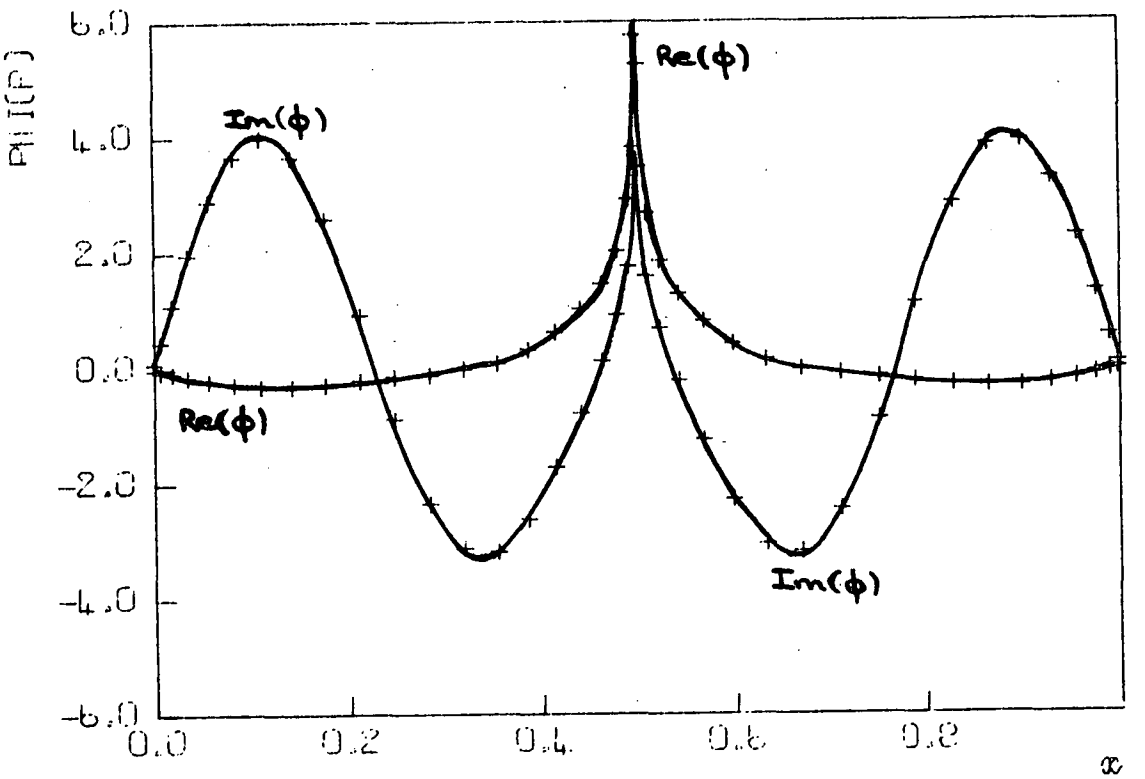


Figure 8. The curves correspond to the same data as those of figure 7, except that the blaze angle of the triangular profile has been increased to 45^{o}.

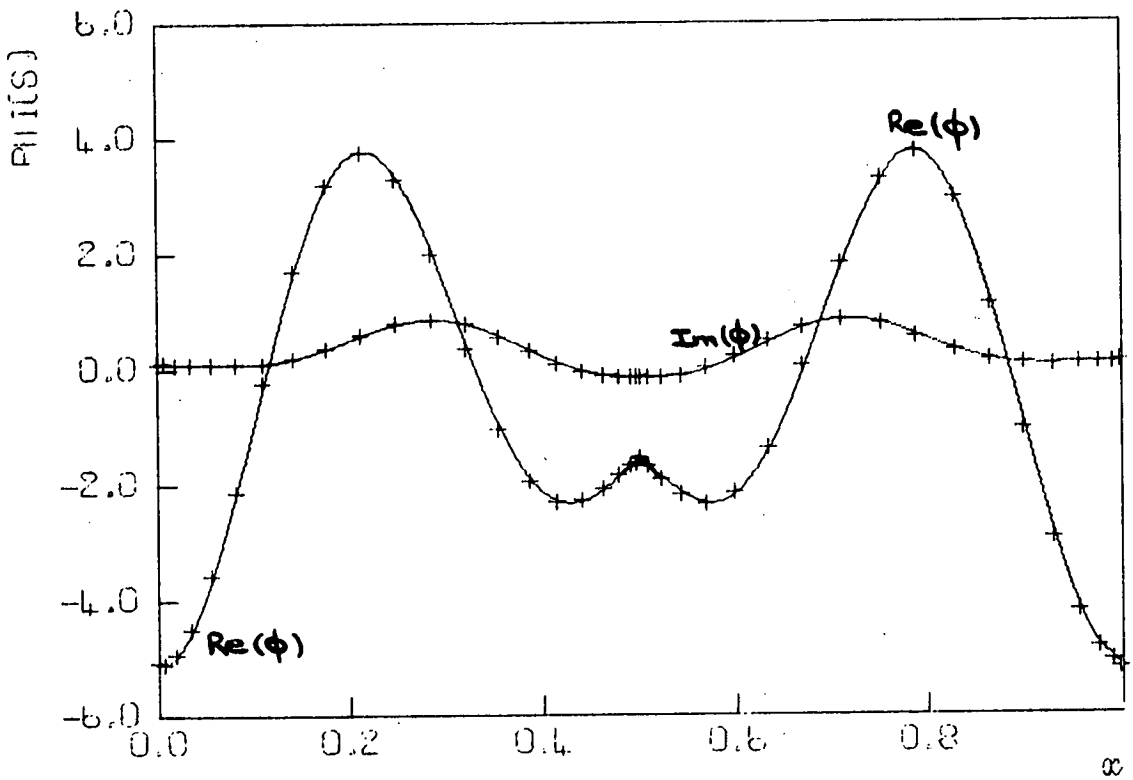


Figure 9. The curves correspond to the same data as those of figure 8, except that the behaviour of the S polarization current density function $\phi(S)(x)$ is shown.

imaginary parts of $\phi(P)$ become much more pronounced as the groove facet angles are increased (i.e., as β is decreased).

In Figure 9, the current density function is plotted for the same data as Figure 8, but now for S polarization. No singularities of $\phi(S)$ can be seen. The concentration of the abscissae of the Gaussian quadrature rule around the segment ends at 0, 0.50 and 1.00 can be seen clearly.

From equation (68), the singularity of $\phi(P)$ at a profile corner is always integrable. It is evident from Figure 8 that the region about the singularity can contribute quite significantly to the area under the curve of $\phi(P)$ (or of $\phi(P)$ multiplied by a harmonic function). Thus, there is a need to have an adequate distribution of integration abscissae in this region. This need can well be satisfied by the use of a Gauss-Legendre quadrature rule, which provides a natural concentration of abscissae at the ends of the interval of integration.

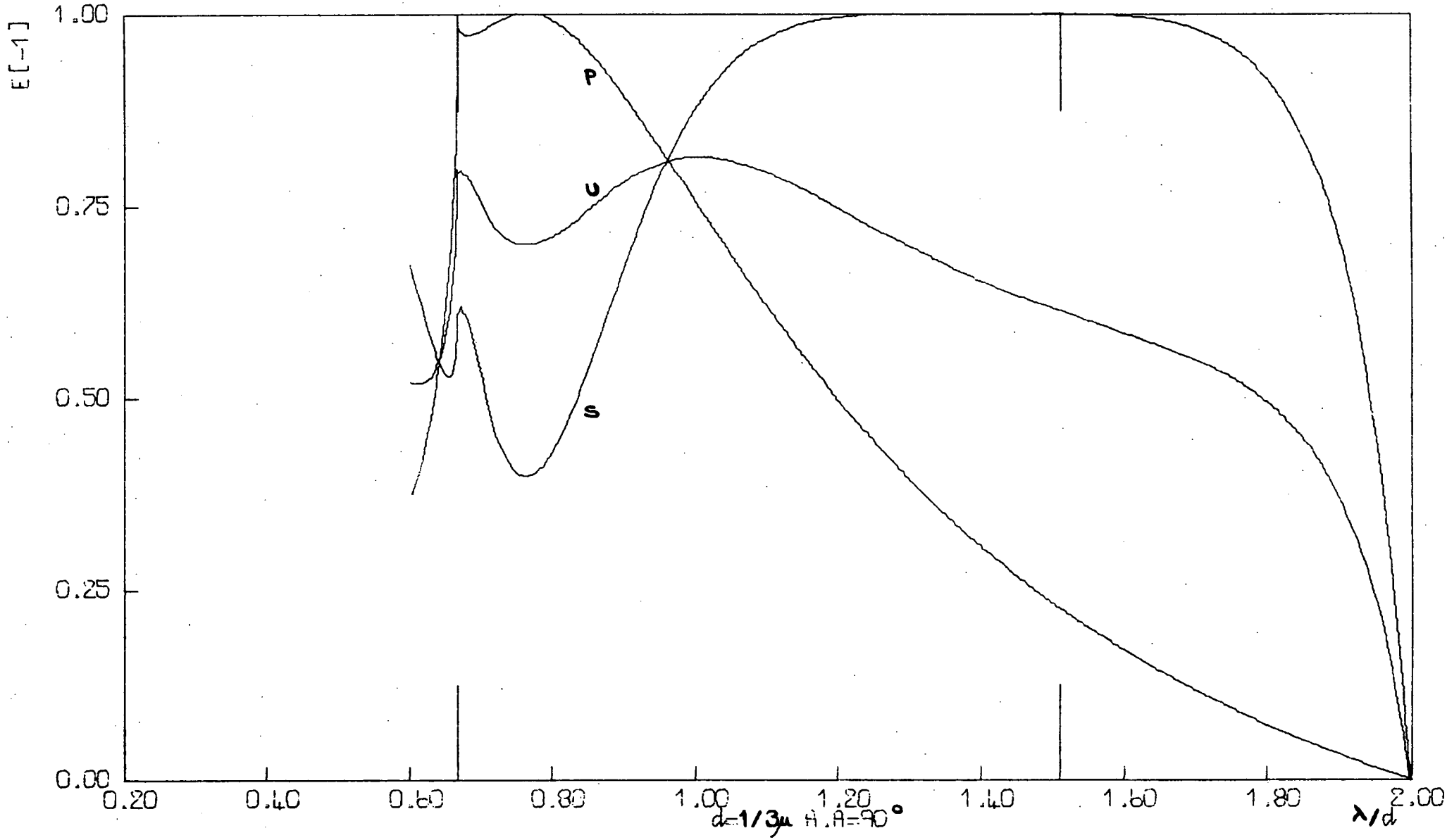
Appendix II

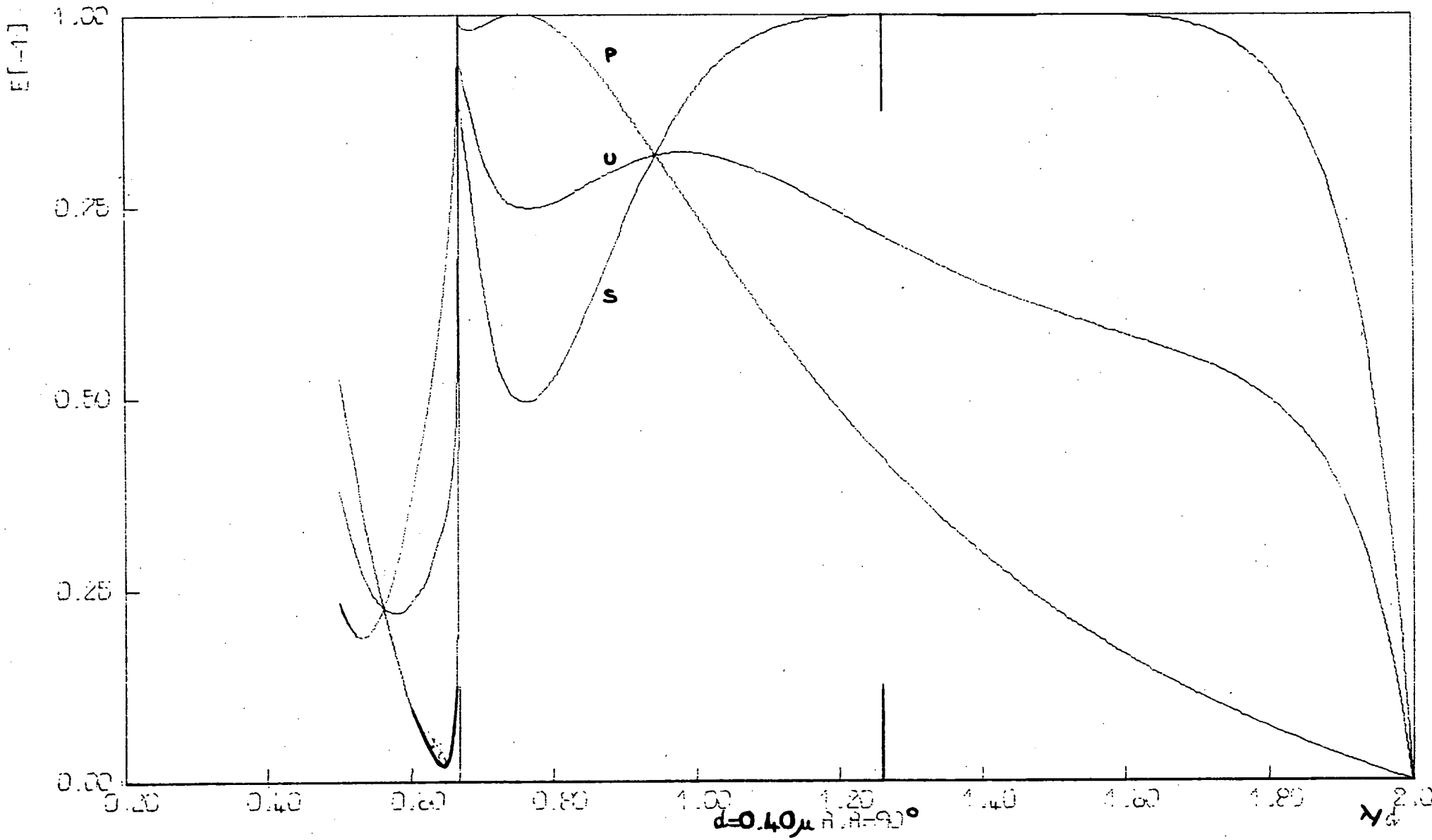
Efficiency curves are given for triangular profile gratings having periods ranging from 1/3 μm to 2.00 μm. The gratings all have groove apex angles of 90^{o}, and are used in a Littrow mount in the order -1. The three curves on each graph show the variation of P, S and U polarization efficiencies with normalized wavelength, for the order -1. The longer bars perpendicular to the wavelength axis indicate the Rayleigh wavelengths, and in addition the geometrical blaze wavelength of the grating. The period (d), the profile angles (α, β) and the blaze wavelength (λ_B) for each grating considered are given in Table 1.

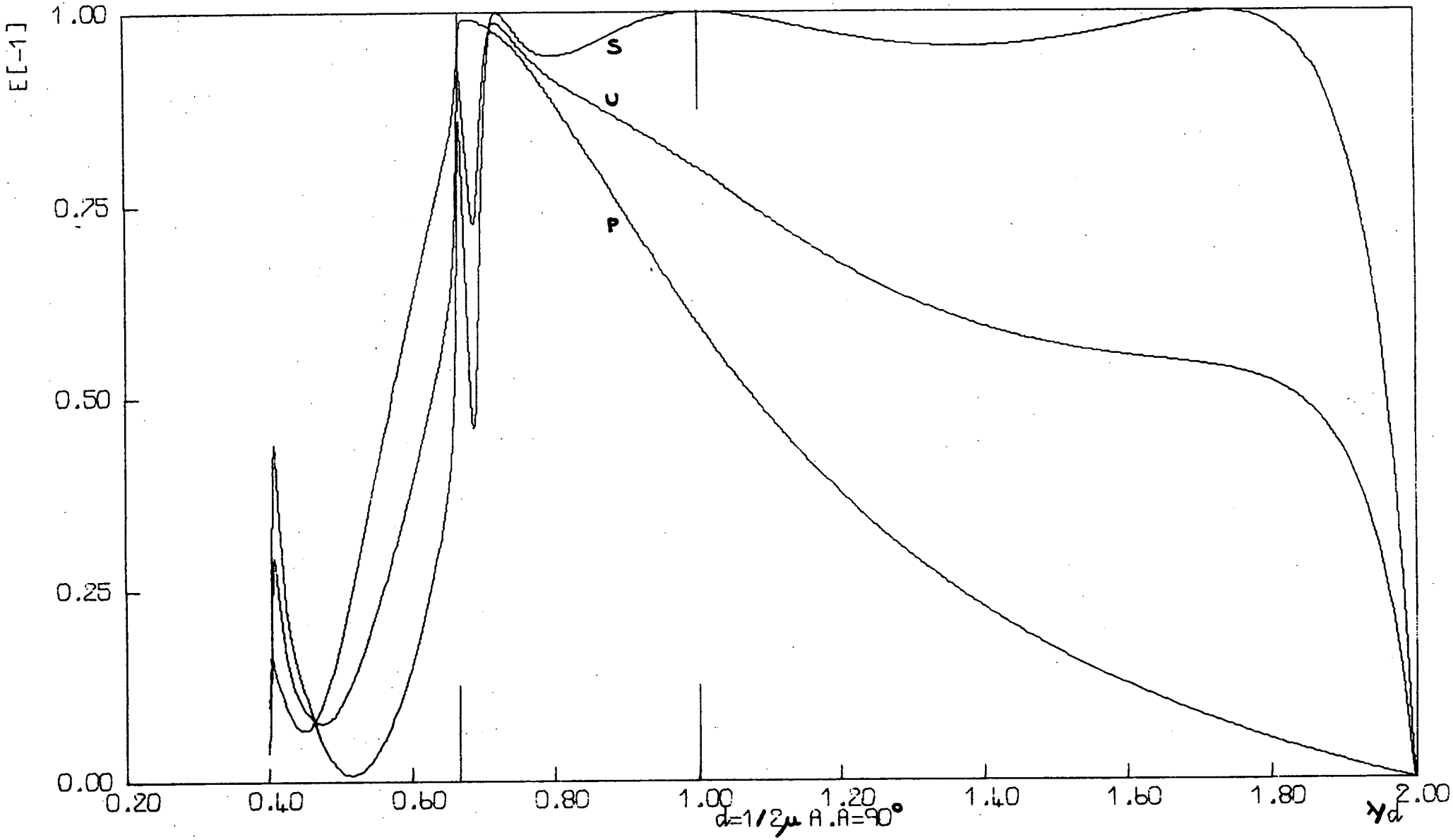
Table 1

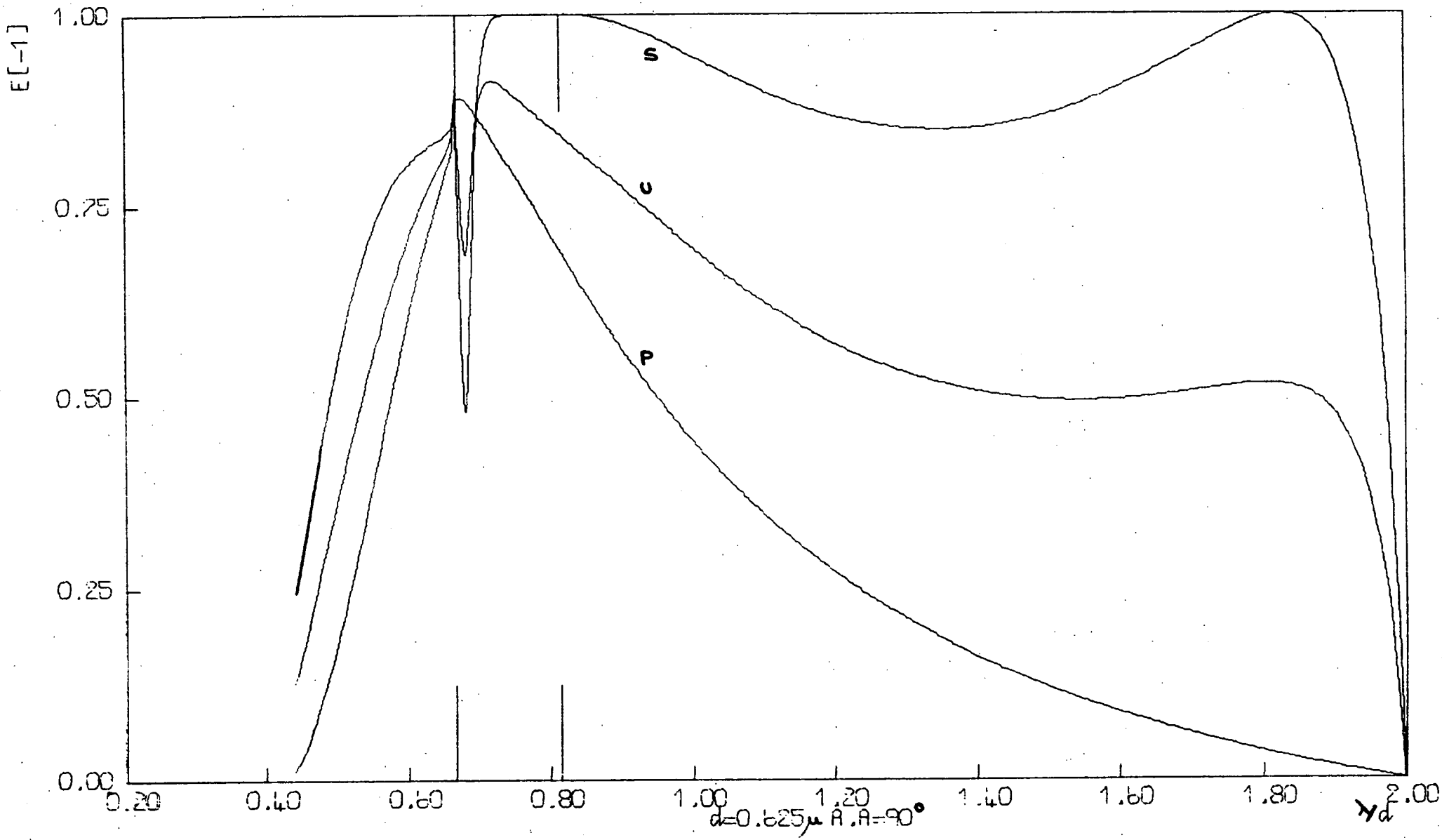
Parameters of the gratings for which efficiency curves are given in this Appendix.

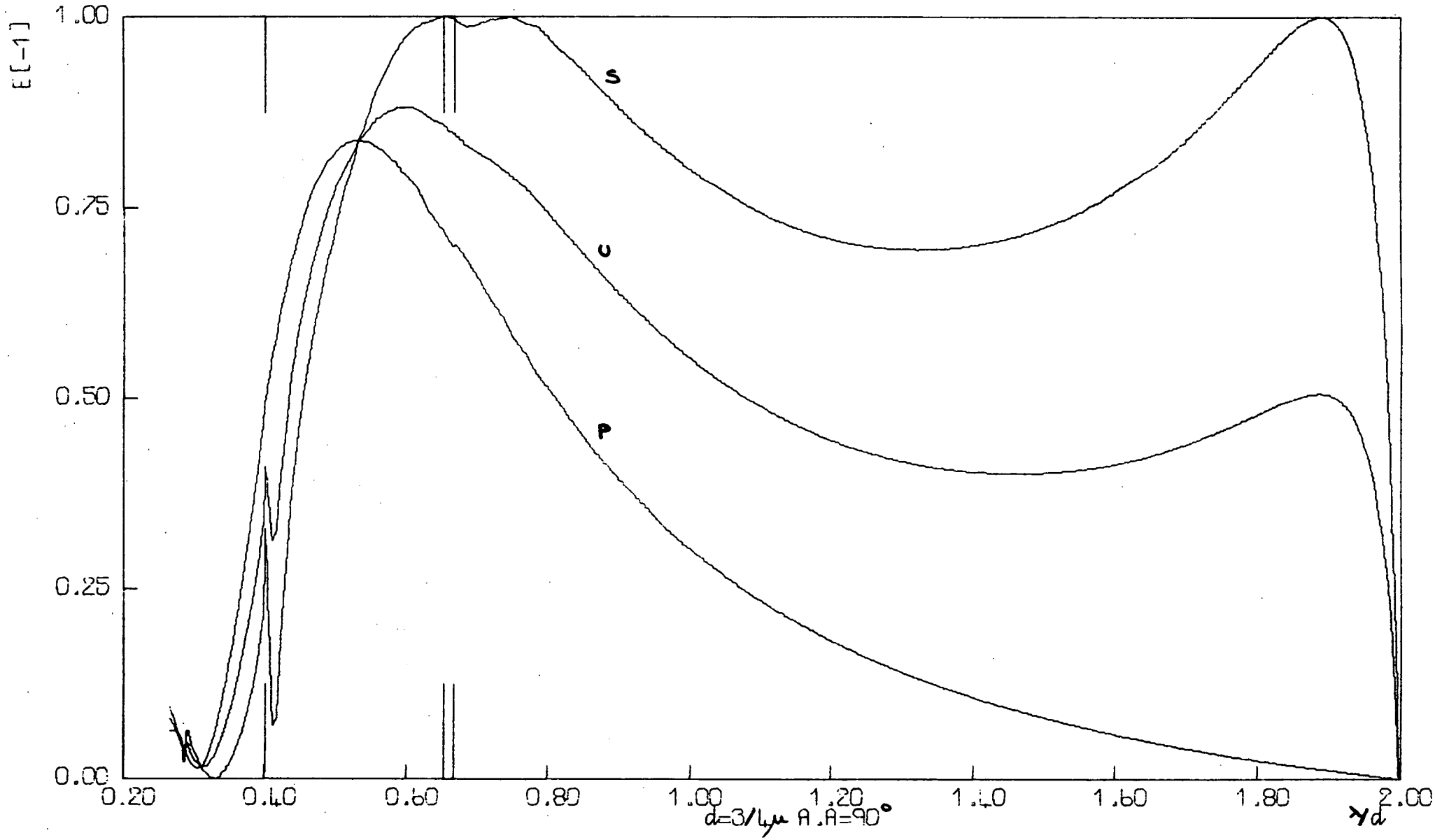
d(μm)	α(°)	β(°)	λ_B(μm)
0.3333	49.0	41.0	0.50309
0.40	39.0	51.0	0.50346
0.50	30.0	60.0	0.50000
0.625	24.0	66.0	0.50842
0.75	19.0	71.0	0.48836
0.875	16.6	73.4	0.49996
1.00	14.5	75.5	0.50076
1.25	11.5	78.5	0.49842
1.50	9.0	81.0	0.46929
2.00	7.2	82.8	0.50133

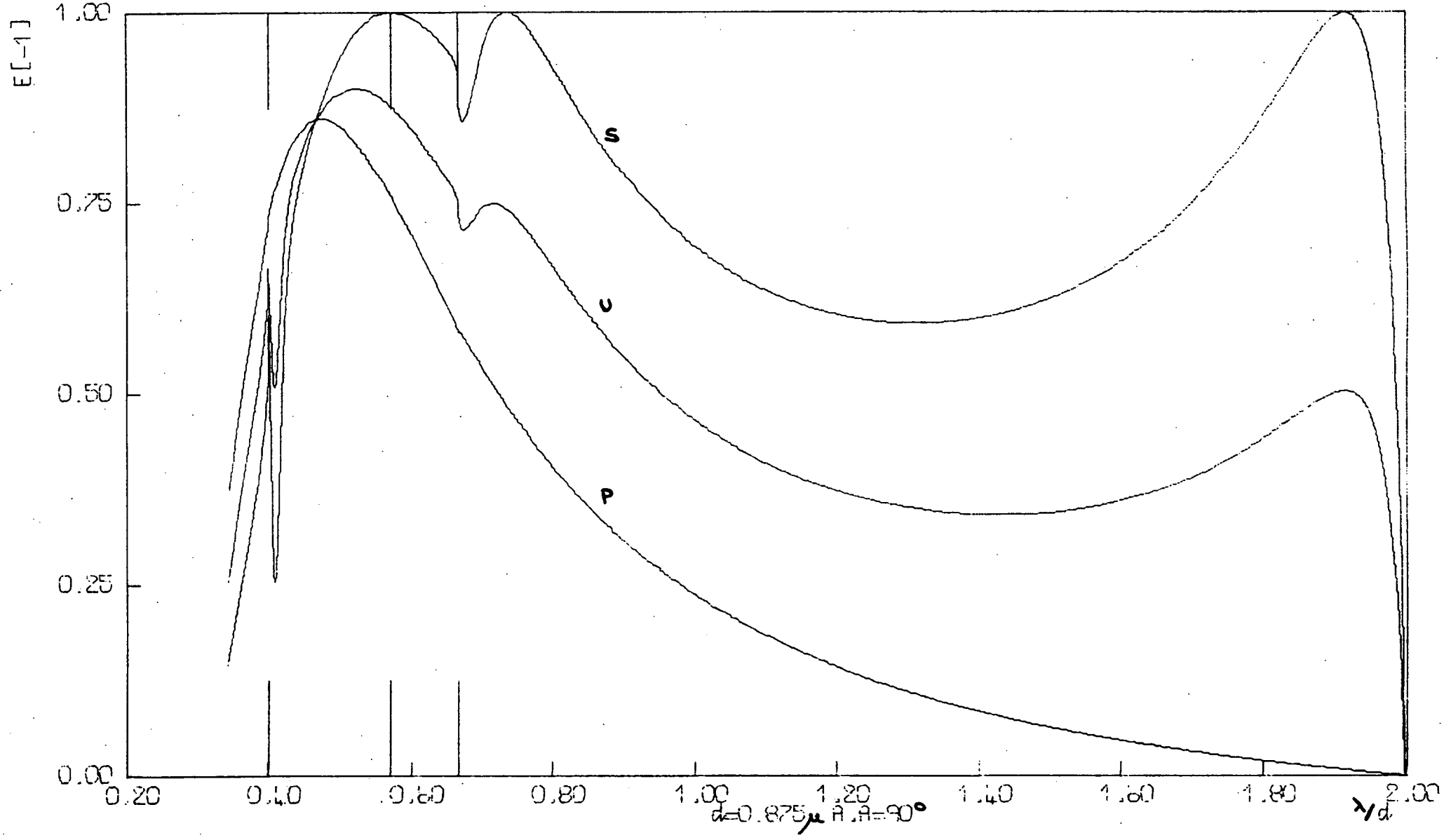

E[-1]
1.00
0.75
0.50
0.25
0.00
P
U
S
0.20
0.40
0.60
0.80
1.00
1.20
1.40
1.60
1.80
2.00
λ/d
d=1/3μ A.A=90°

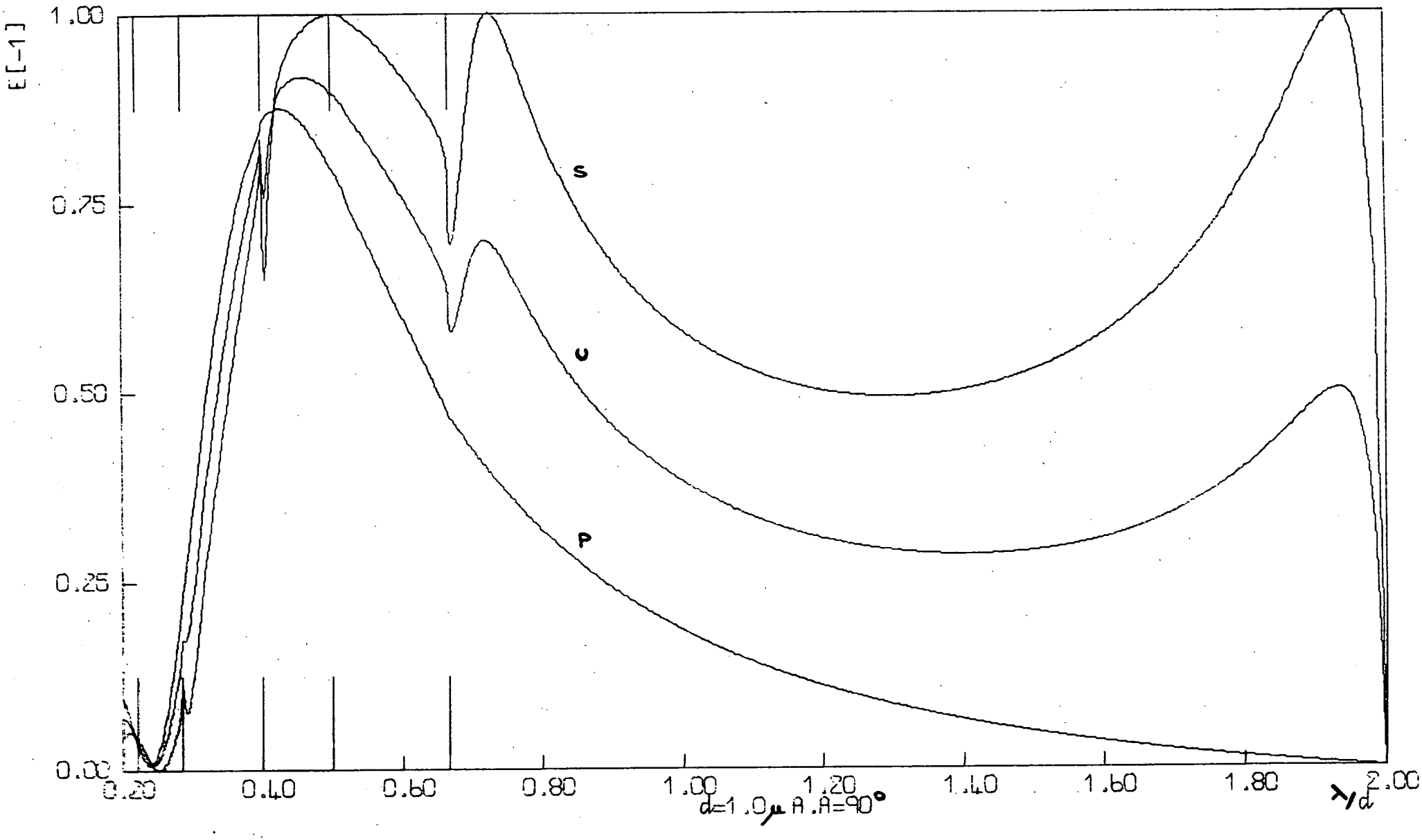

R[-]
1.00
0.75
0.50
0.25
0.00
P
U
S
0.20
0.40
0.60
0.80
1.00
1.20
1.40
1.60
1.80
2.00
λ/d
d=0.40μ m A=90°

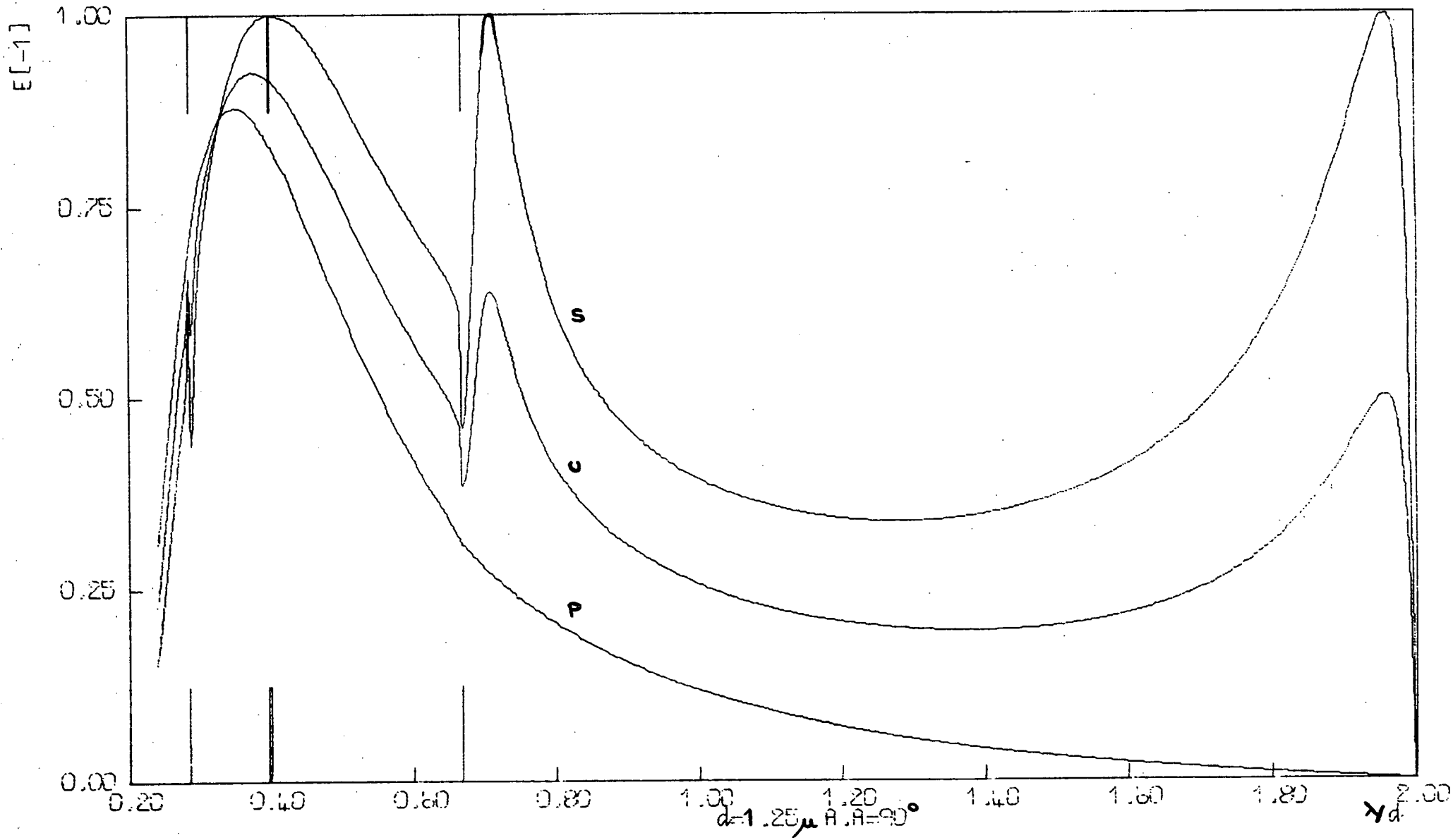

E[-1]
1.00
0.75
0.50
0.25
0.00
S
U
P
0.20
0.40
0.60
0.80
1.00
1.20
1.40
1.60
1.80
2.00
d=1/2μ A.A=90°
γd

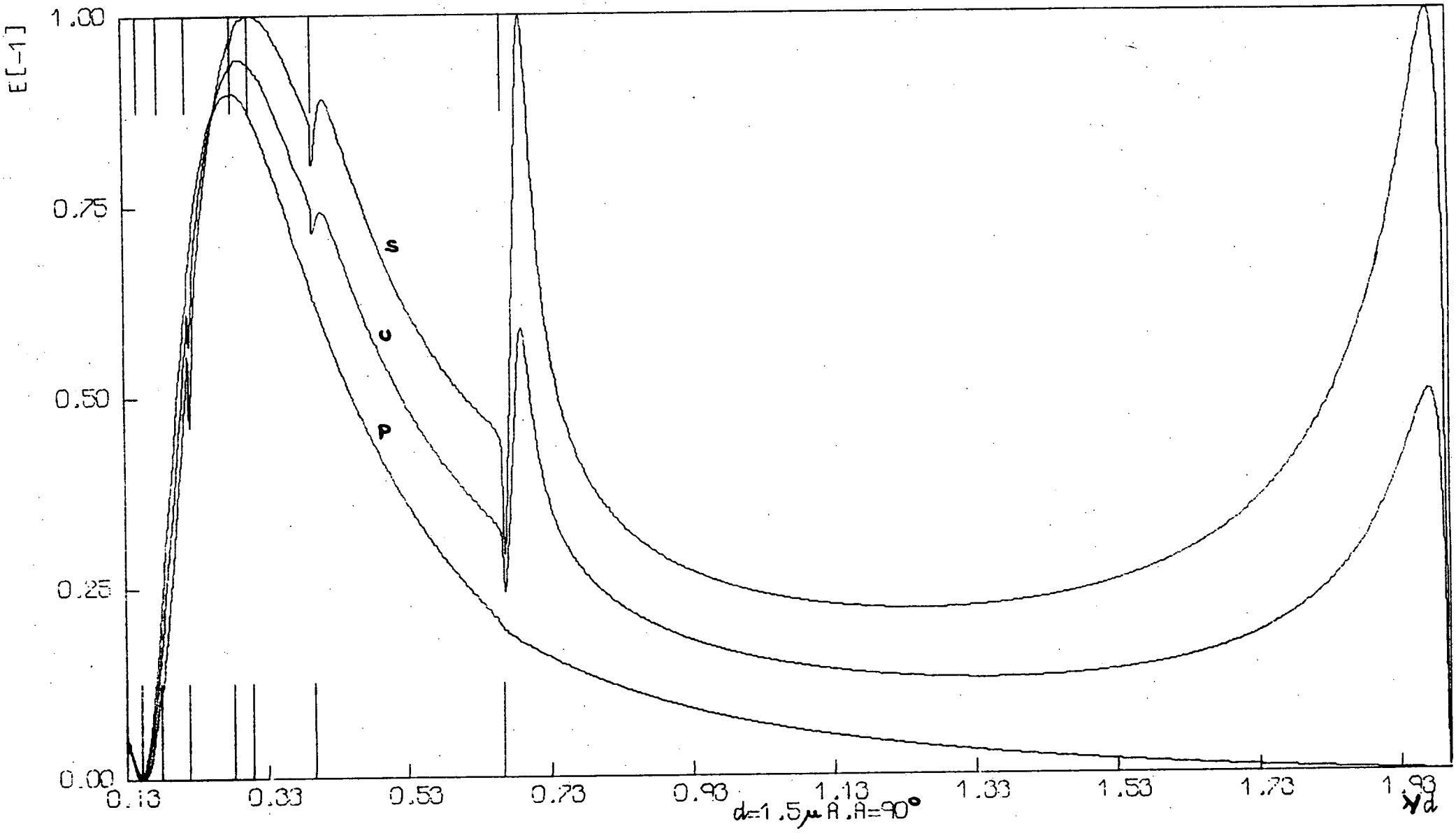

E[-1]
1.00
0.75
0.50
0.25
0.00
S
U
P
0.20
0.40
0.60
0.80
1.00
1.20
1.40
1.60
1.80
2.00
d=0.625μ Å.A=90°
γd

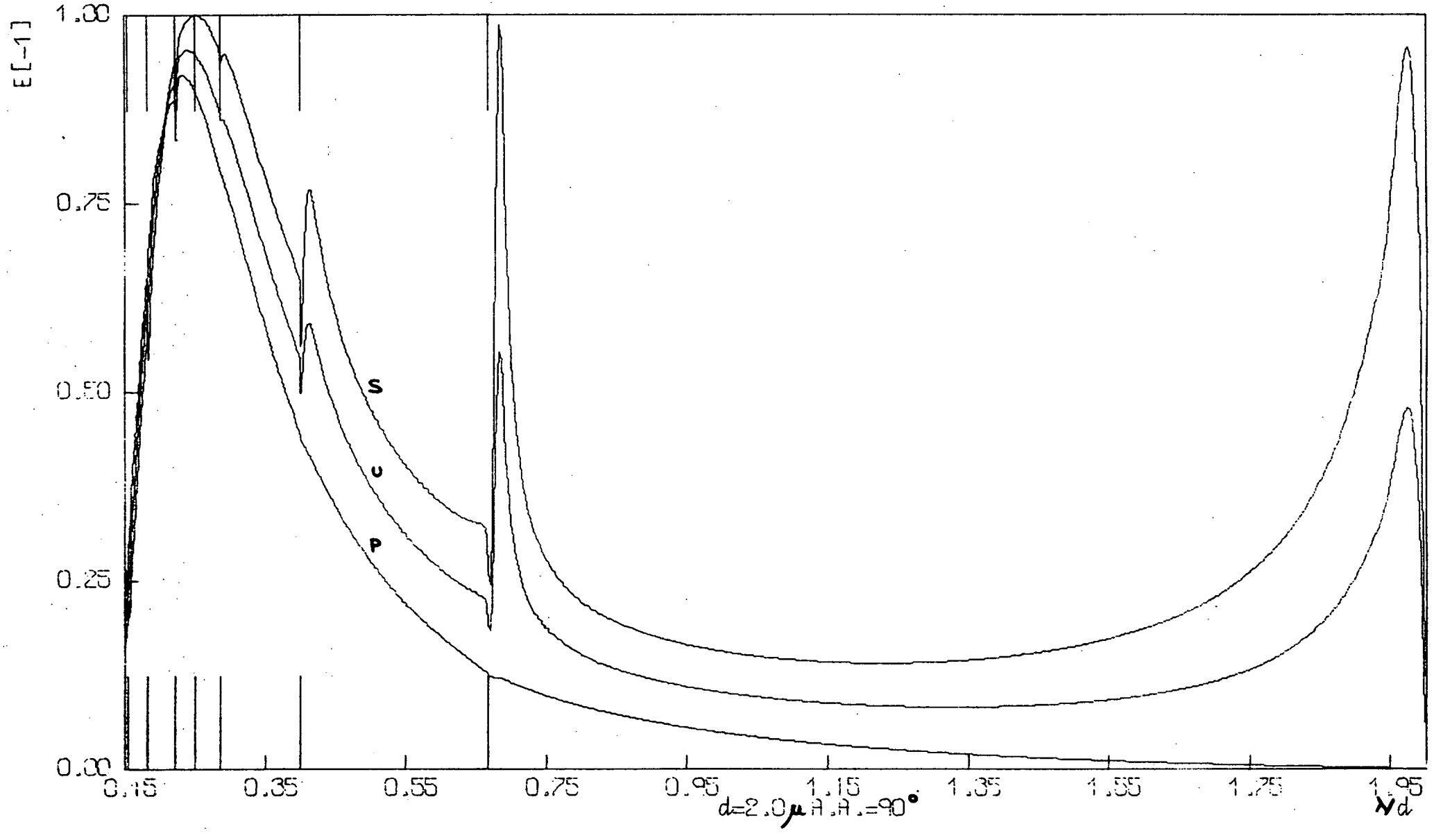

E[-1]
1.00
0.75
0.50
0.25
0.00
S
C
P
0.20
0.40
0.60
0.80
1.00
1.20
1.40
1.60
1.80
2.00
d=3/4μ A.A=90°
x/d


E[-1]
1.00
0.75
0.50
0.25
0.00
S
U
P
0.20
0.40
0.60
0.80
1.00
1.20
1.40
1.60
1.80
2.00
λ/d
d=0.875μ A.A=90°


E[-1]
1.00
0.75
0.50
0.25
0.00
S
U
P
0.20
0.40
0.60
0.80
1.00
1.20
1.40
1.60
1.80
2.00
λ/d
d=1.0μ A.A=90°

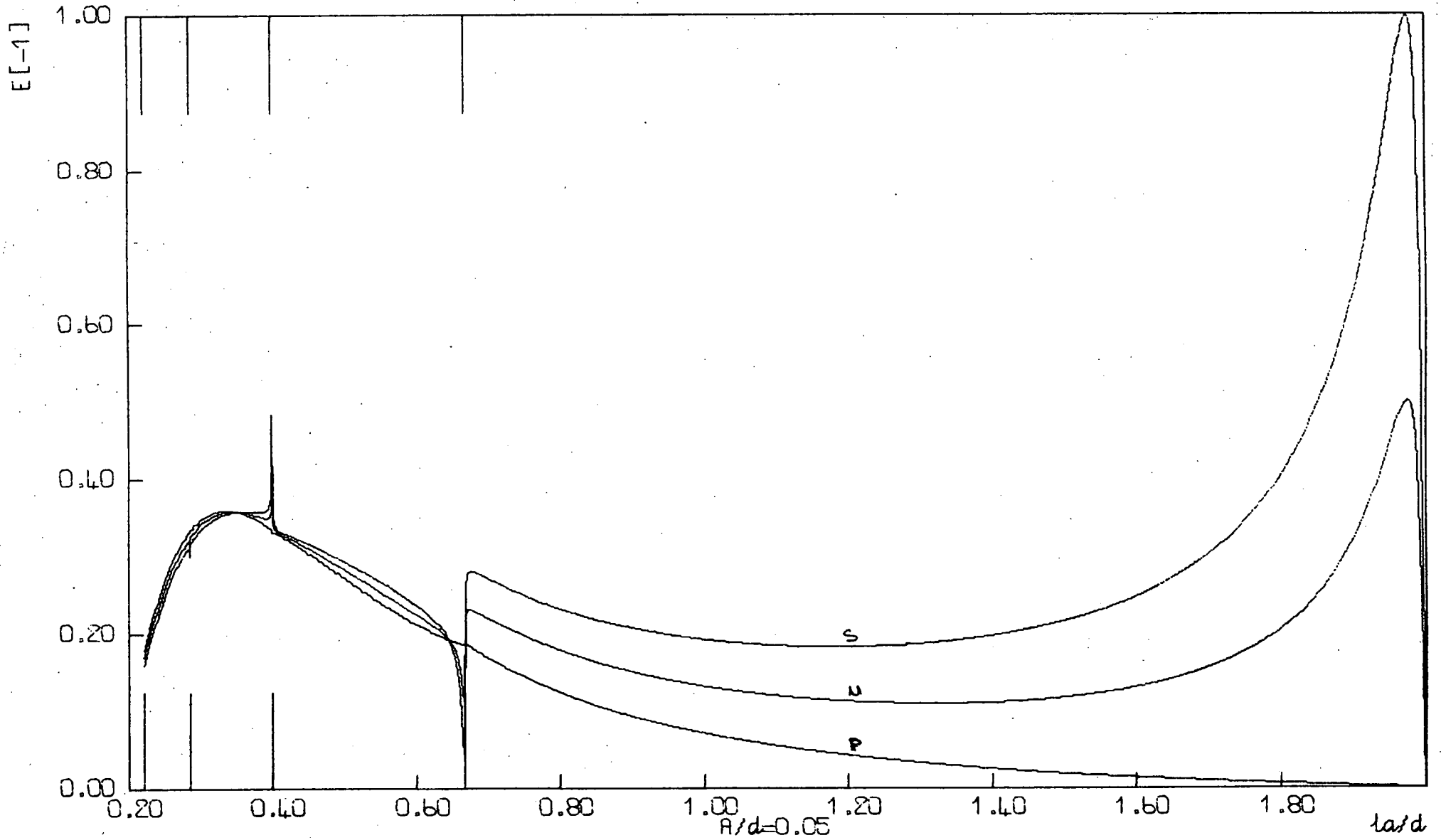

E[-1]
1.00
0.75
0.50
0.25
0.00
0.20
0.40
0.60
0.80
1.00
1.20
1.40
1.60
1.80
2.00
S
C
P
d=1.25μ A.A=90°
λd

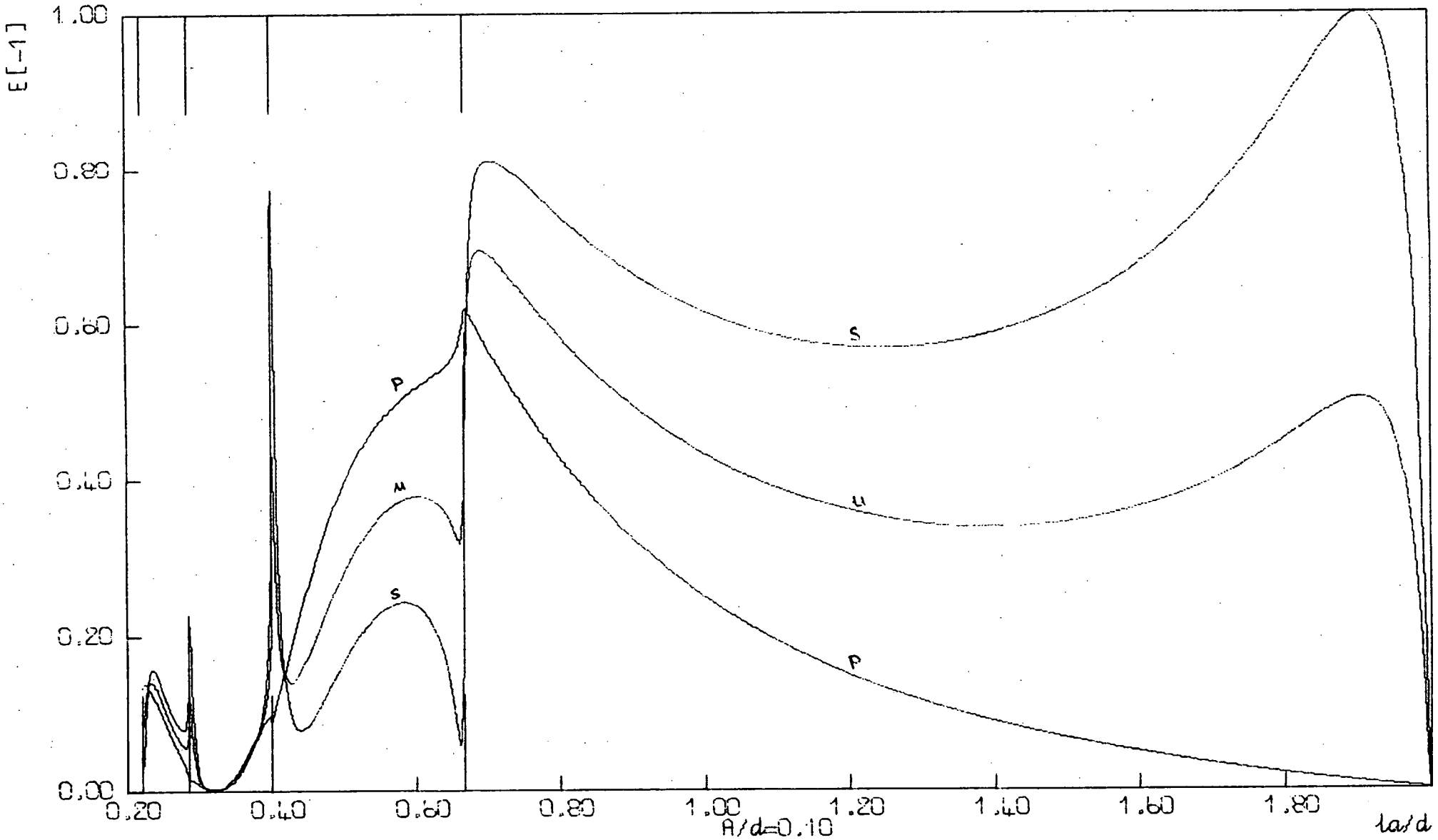

E[-1]
1.00
0.75
0.50
0.25
0.00
S
U
P
0.13
0.33
0.53
0.73
0.93
1.13
1.33
1.53
1.73
1.93
d=1.5μm Å.Å=90°
x/d

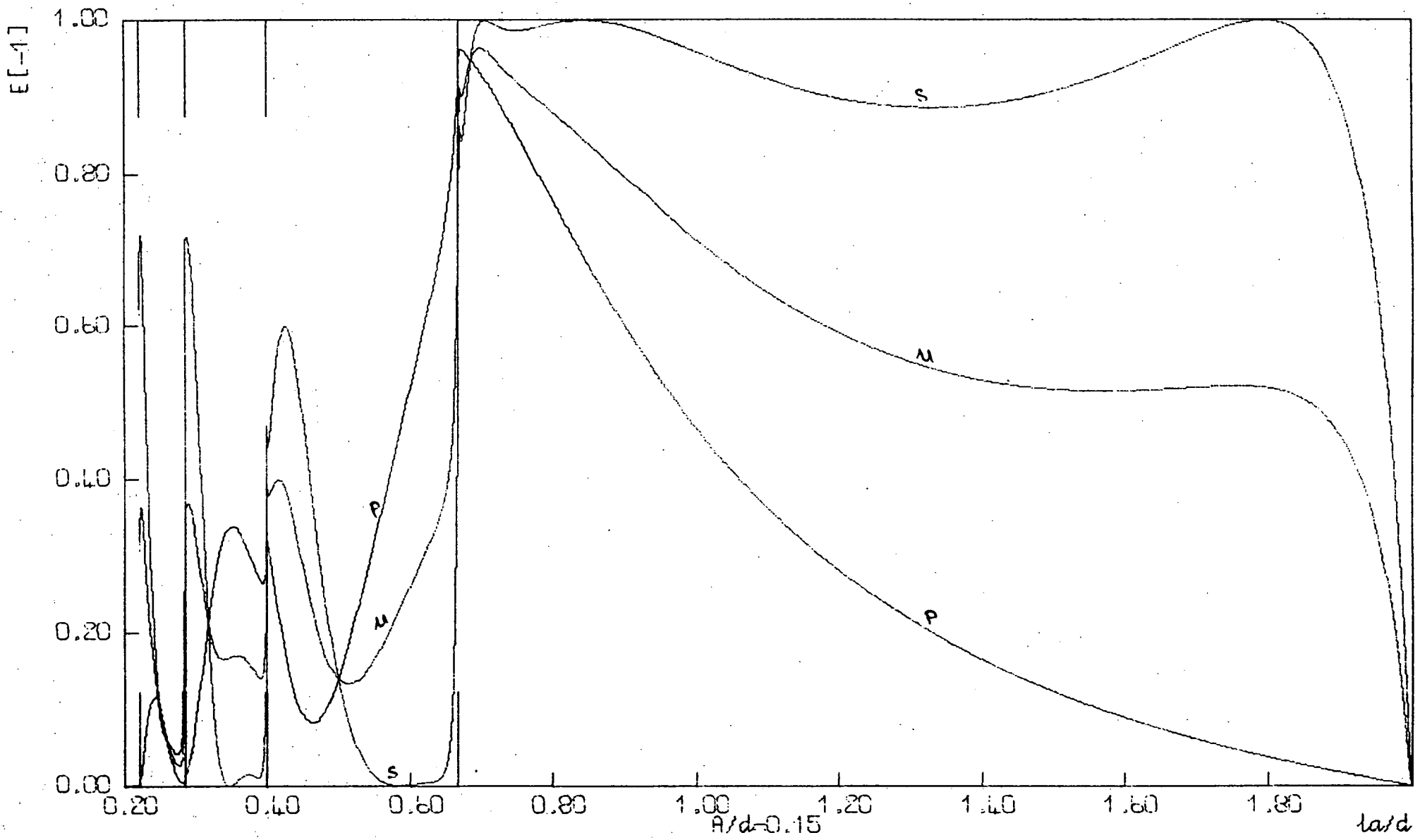

E[-1]
1.00
0.75
0.50
0.25
0.00
S
U
P
0.15
0.35
0.55
0.75
0.95
1.15
1.35
1.55
1.75
1.95
d=2.0μ A.A.=90°
νd

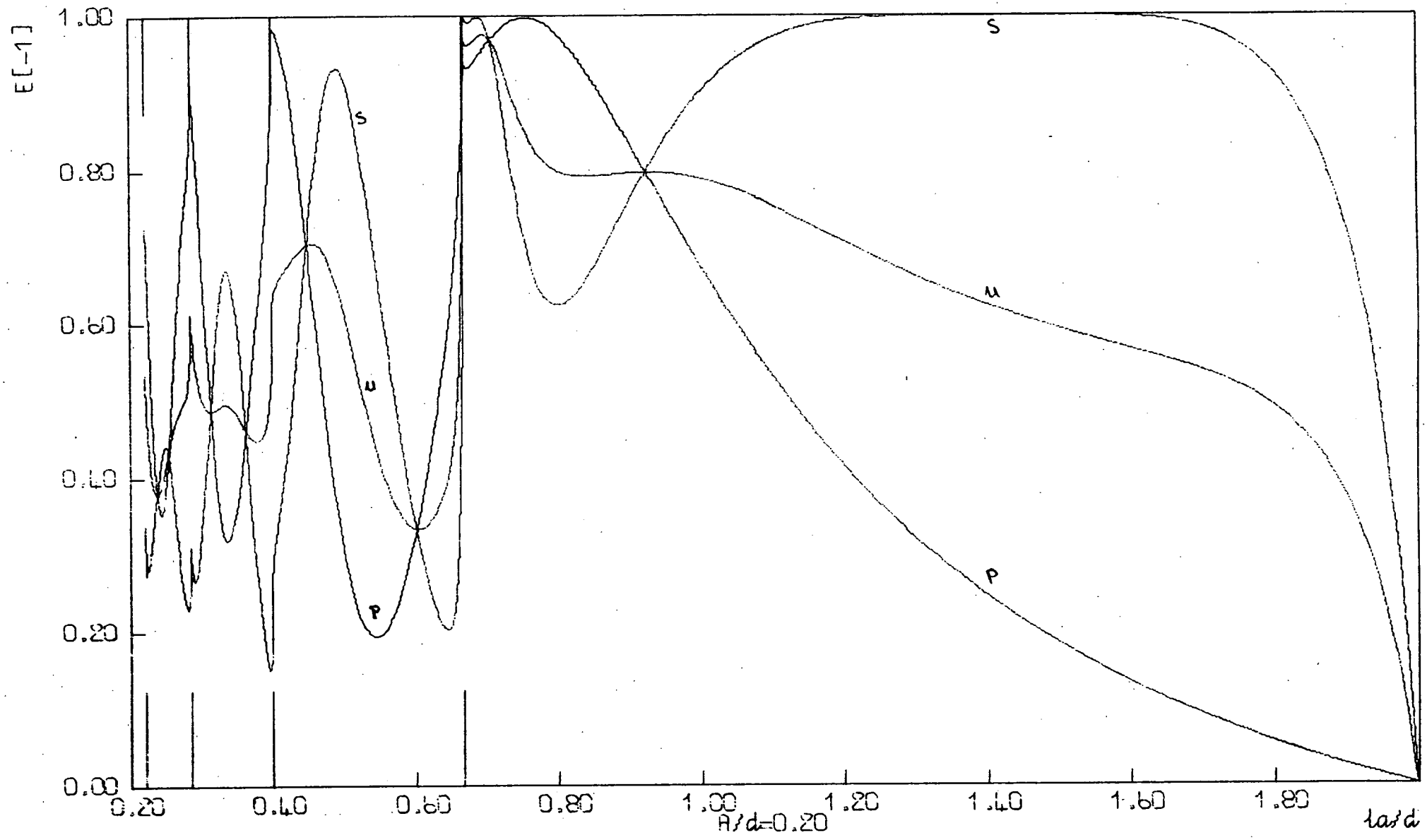
Appendix III

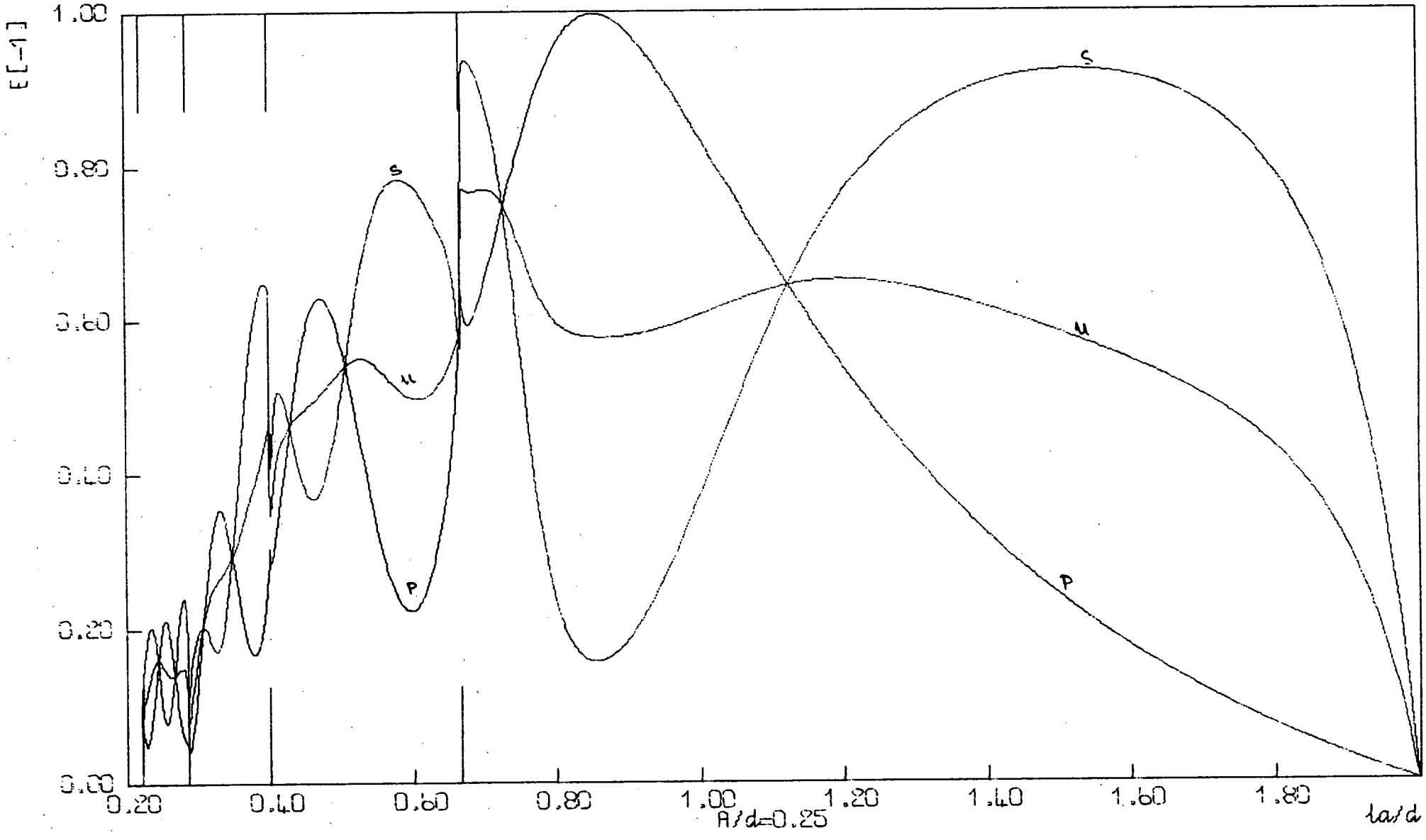
The set of curves given in this Appendix are for a sinusoidal profile grating used in a Littrow mount in the order -1. The three curves on each graph give the variation of the efficiency in order -1 with normalized wavelength for P, S and U polarized light. The Rayleigh wavelengths are indicated by the bars perpendicular to the normalized wavelength axis. The normalized profile amplitude is labelled under each graph. (These efficiency curves have also been given in reference (6.10).)

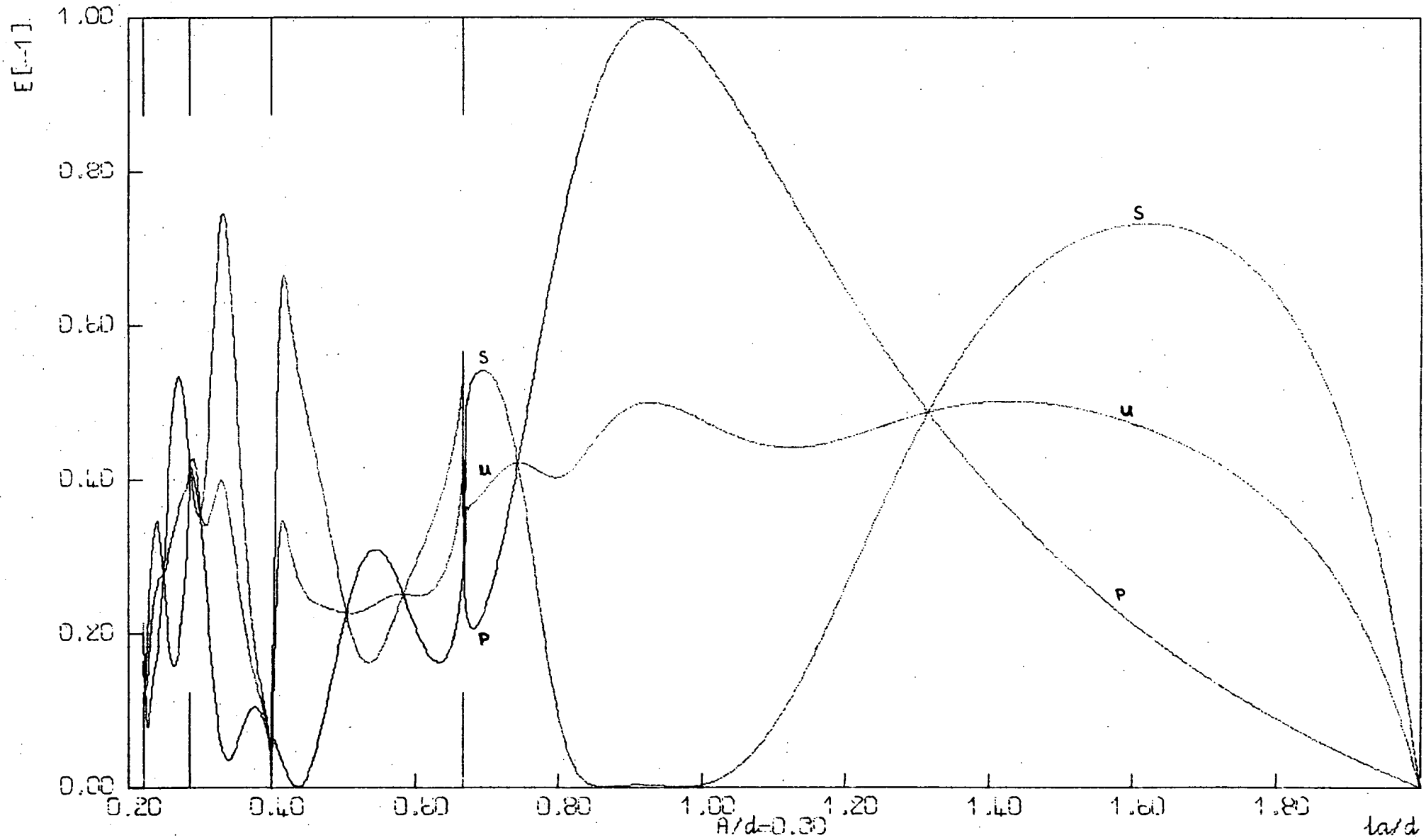

E[-1]
1.00
0.80
0.60
0.40
0.20
0.00
0.20
0.40
0.60
0.80
1.00
1.20
1.40
1.60
1.80
a/d
A/d=0.05
S
U
P

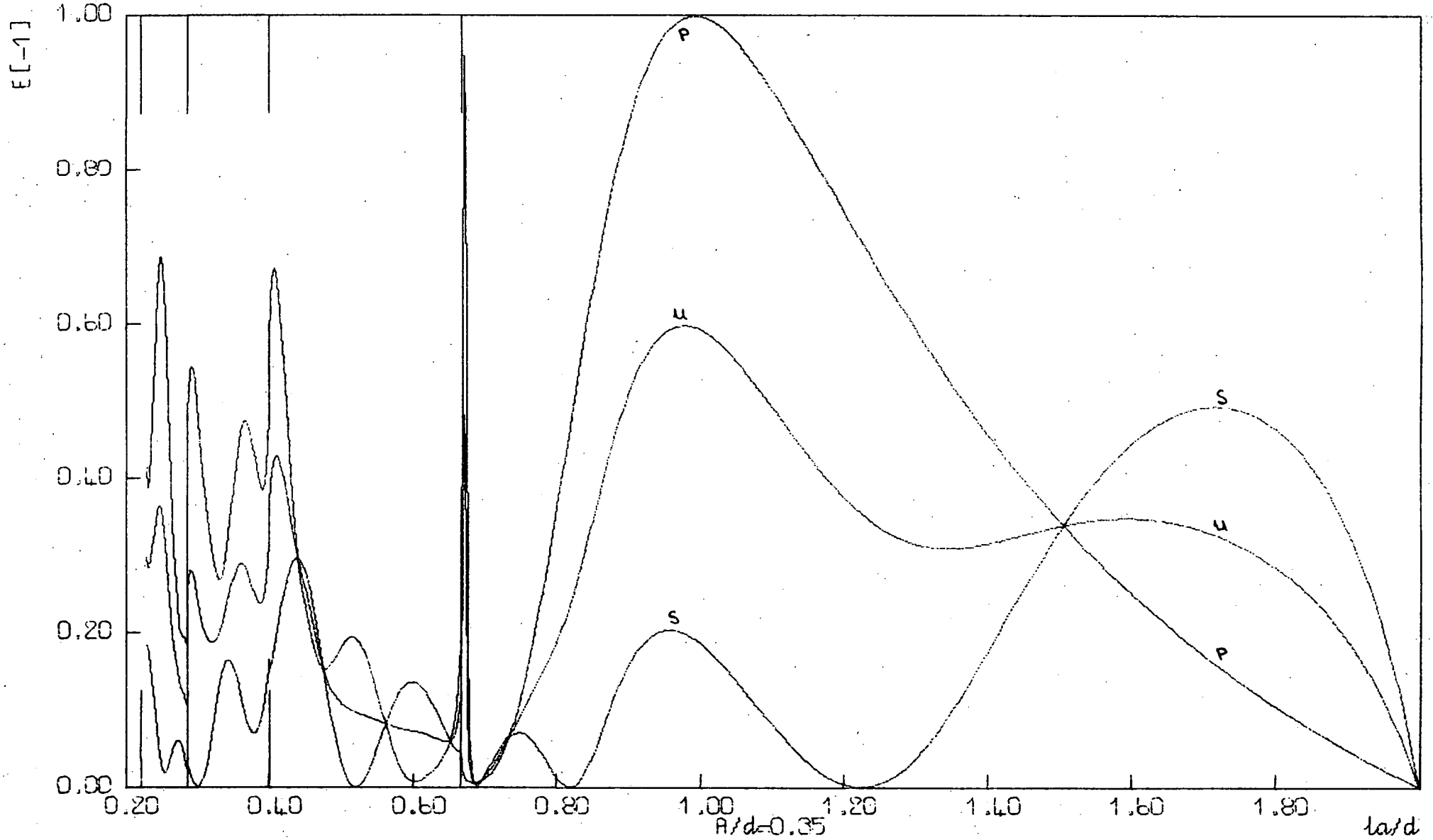

E[-1]
1.00
0.80
0.60
0.40
0.20
0.00
0.20
0.40
0.60
0.80
1.00
1.20
1.40
1.60
1.80
A/d=0.10
la/d
P
M
S
S
U
P

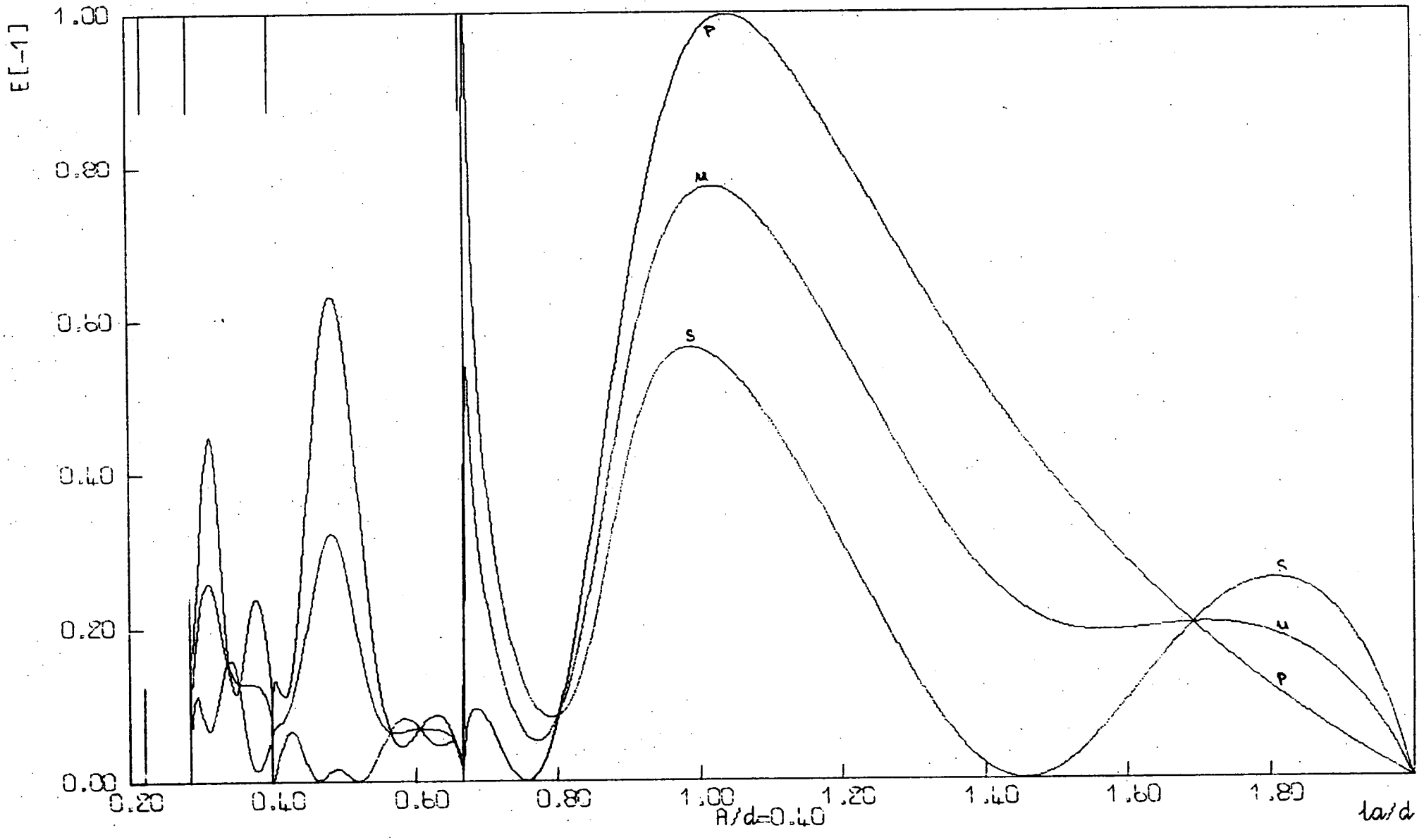

E[-1]
1.00
0.80
0.60
0.40
0.20
0.00
0.20
0.40
0.60
0.80
1.00
1.20
1.40
1.60
1.80
A/d=0.15
la/d
S
M
P
P
M
S

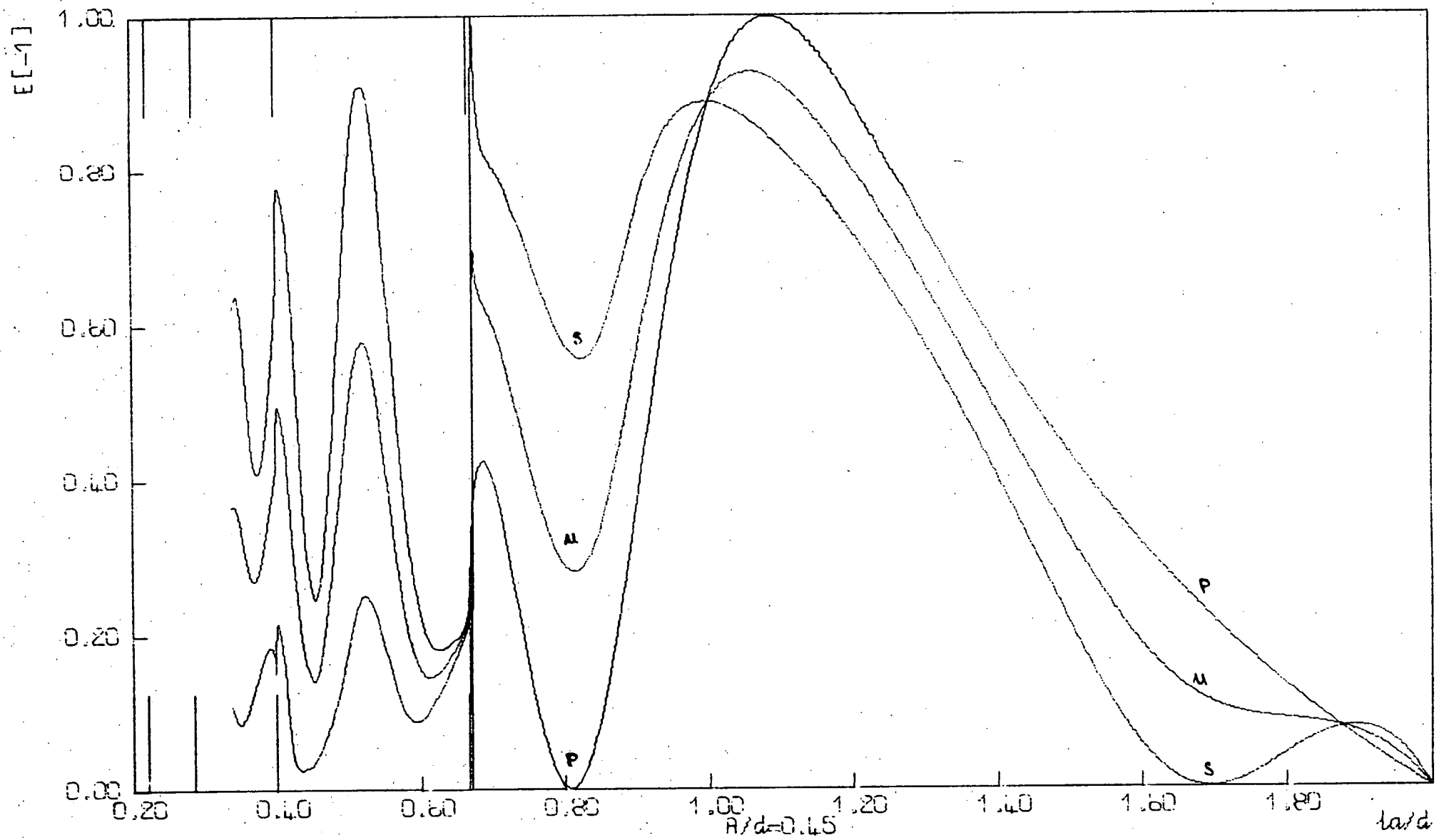

E[-1]
1.00
0.80
0.60
0.40
0.20
0.00
0.20
0.40
0.60
0.80
1.00
1.20
1.40
1.60
1.80
las/d
A/d=0.20
S
U
P
S
U
P

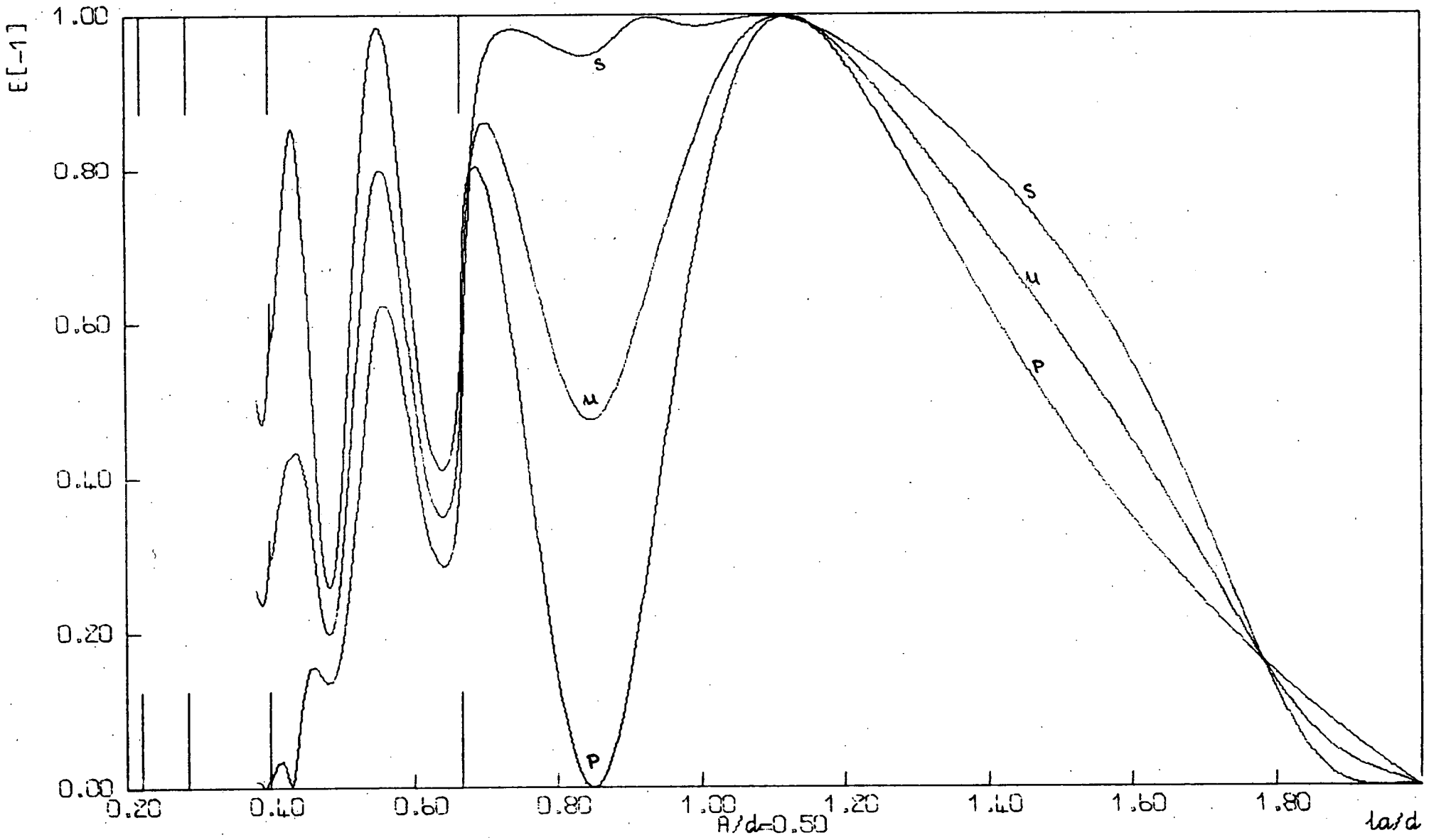

E[-1]
1.00
0.80
0.60
0.40
0.20
0.00
0.20
0.40
0.60
0.80
1.00
1.20
1.40
1.60
1.80
A/d=0.25
la/d
S
U
P
S
U
P

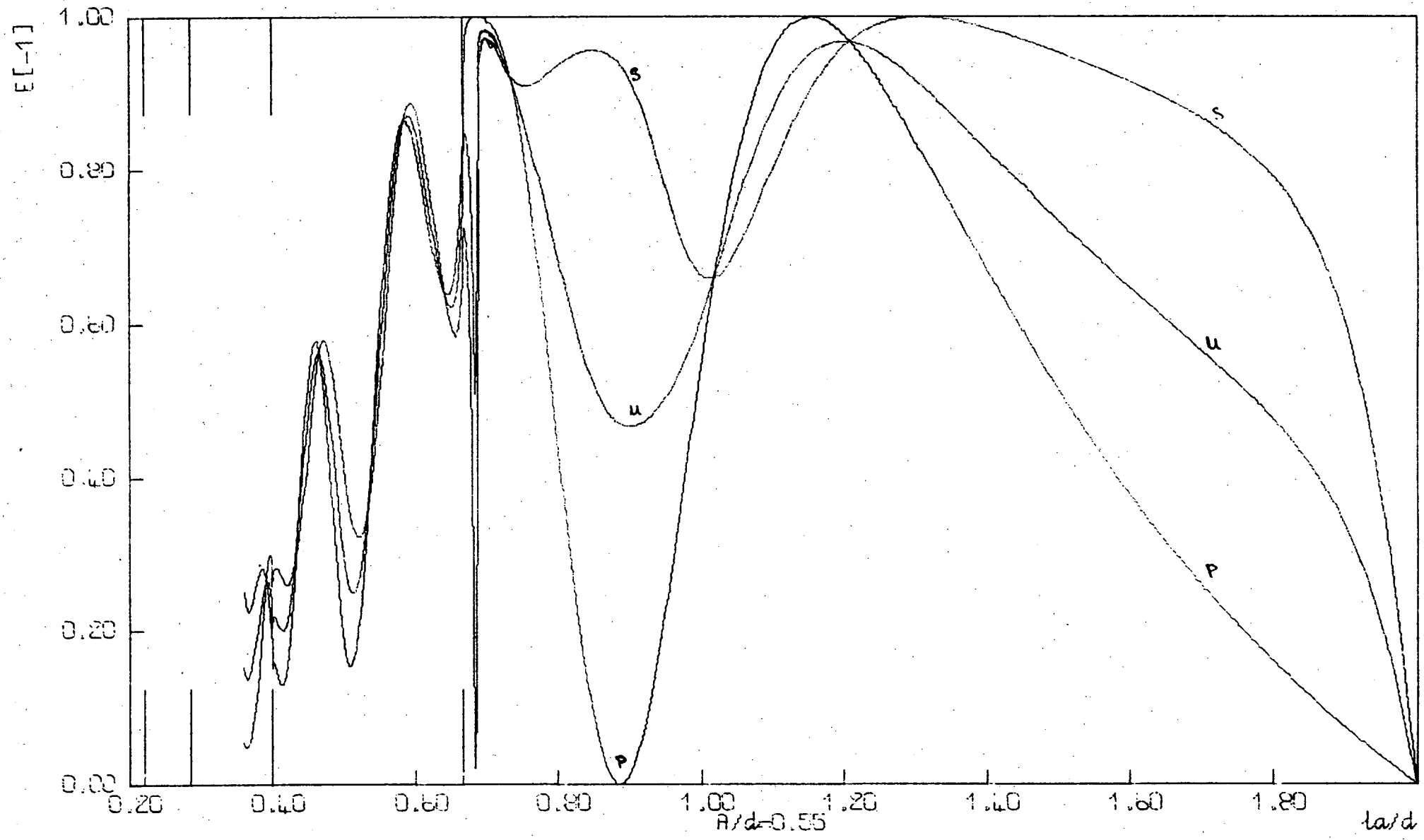

E[-1]
1.00
0.80
0.60
0.40
0.20
0.00
0.20
0.40
0.60
0.80
1.00
1.20
1.40
1.60
1.80
A/d=0.30
t a/d
S
u
P
S
u
P

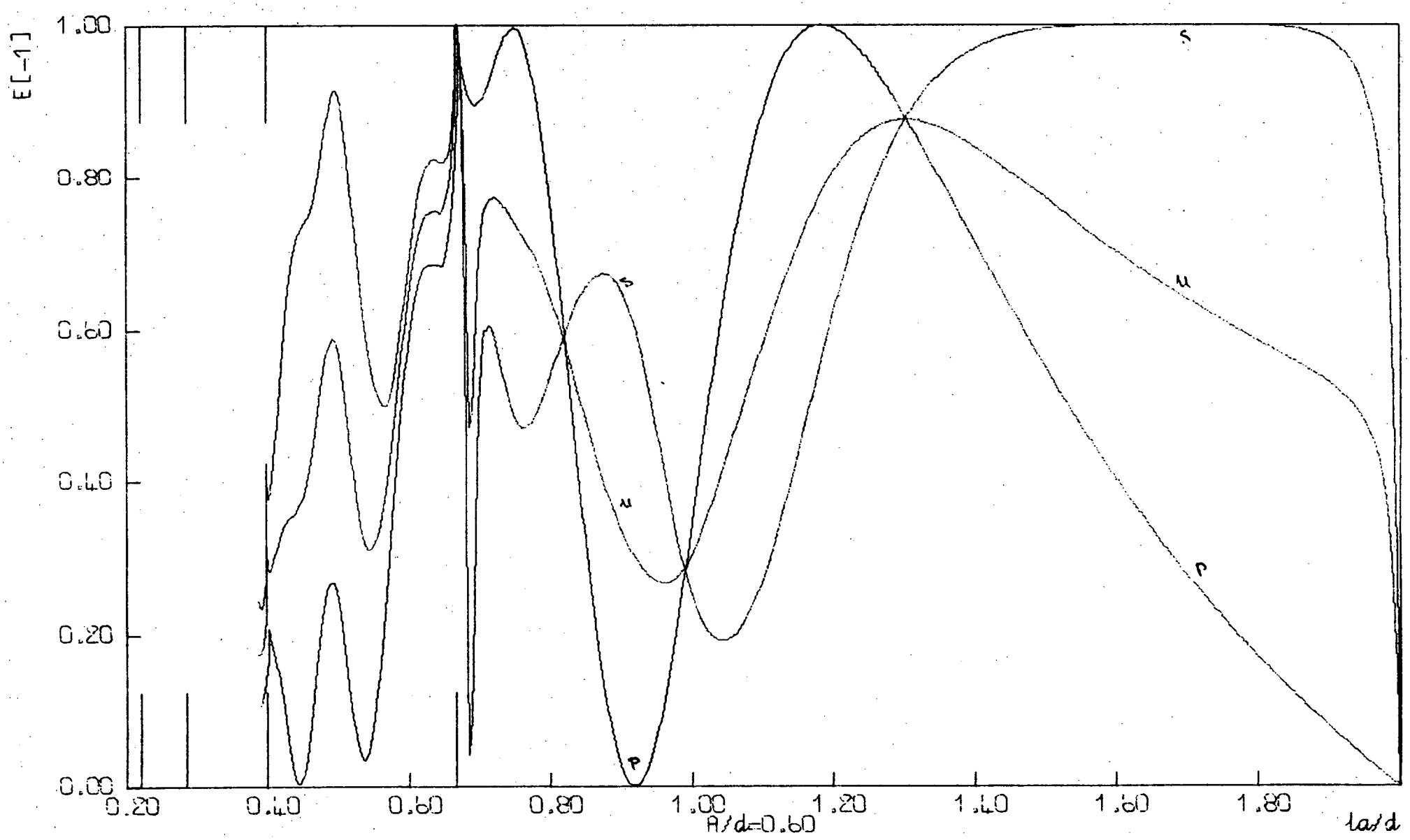

E[-1]
1.00
0.80
0.60
0.40
0.20
0.00
0.20
0.40
0.60
0.80
1.00
1.20
1.40
1.60
1.80
λ/d
A/d=0.35
P
U
S
S
U
P

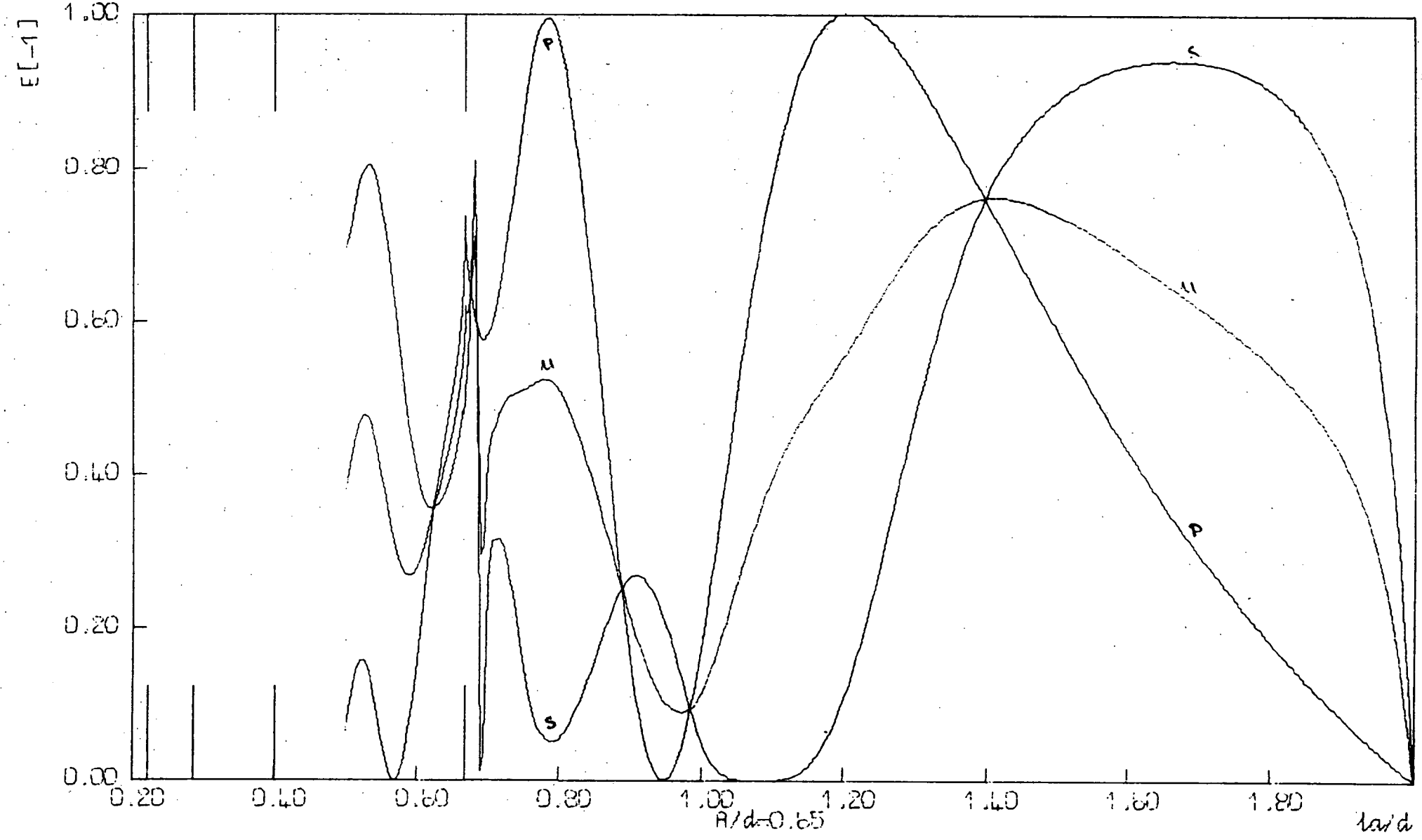

E[-1]
1.00
0.80
0.60
0.40
0.20
0.00
P
M
S
S
M
P
0.20
0.40
0.60
0.80
1.00
1.20
1.40
1.60
1.80
R/d=0.40
la/d


E[-1]
1.00
0.80
0.60
0.40
0.20
0.00
0.20
0.40
0.60
0.80
1.00
1.20
1.40
1.60
1.80
λ/d
A/d=0.45
S
M
P
P
M
S


E[-1]
1.00
0.80
0.60
0.40
0.20
0.00
0.20
0.40
0.60
0.80
1.00
1.20
1.40
1.60
1.80
λ/d
A/d=0.50
S
M
P
S
M
P


E[-1]
1.00
0.80
0.60
0.40
0.20
0.00
0.20
0.40
0.60
0.80
1.00
1.20
1.40
1.60
1.80
λ/d=0.55
λ/d
S
U
P
S
U
P


E[-1]
1.00
0.80
0.60
0.40
0.20
0.00
0.20
0.40
0.60
0.80
1.00
1.20
1.40
1.60
1.80
A/d=0.60
la/d
S
U
P
Z
P


E[-1]
1.00
0.80
0.60
0.40
0.20
0.00
0.20
0.40
0.60
0.80
1.00
1.20
1.40
1.60
1.80
λ/d=0.65
λ₀/d
P
M
S
S
M
P